Conservation of the SOS operon, *umuDC*, in *Acinetobacter* species

---

A Thesis

Presented to

the Faculty of the College of Science & Technology

Morehead State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

---

by

Ching-Li Lin

May 01, 2007

Accepted by the faculty of the College of Science and Technology, Morehead State University, in partial fulfillment of the requirements for the Master of Science degree.

Janelle Hare
Director of Thesis

Master's Committee: Janelle Hare **, Chair**

David Taylor

May 1, 2007
Date

Conservation of the SOS operon, *umuDC*, in *Acinetobacter* species

Ching-Li Lin, M.S.
Morehead State University, 2007

Director of Thesis: Janelle Hare

This research investigates potential DNA damage response mechanisms in the ubiquitous gram-negative bacterial genus, *Acinetobacter*. Many bacteria, such as *Escherichia coli*, use an inducible genetic system to respond to DNA damage (the SOS response). For example, products of the *umuDC* SOS operon, encoding DNA polymerase V, carries out error-prone translesion DNA synthesis after DNA damage (SOS mutagenesis). The SOS response requires RecA-mediated self-cleavage of (i) LexA repressor to induce the SOS response genes and (ii) UmuD homodimers to form DNA Pol V. Several proteases such as Lon and ClpXP also regulate the level of UmuD and UmuD' in cells, respectively. Interestingly, in one *Acinetobacter* species, *A. baylyi* strain ADP1, the *umuD* gene is 1.5-fold longer than other *umuD* genes due to an extended 5' region, and the *umuC* gene is mutated. This project examines whether the *umuDC* genes from diverse *Acinetobacter* strains have an atypical *umuD* and a truncated *umuC* gene like ADP1, which could indicate potential abilities to respond to DNA damage. Thus, without both UmuD and UmuC, ADP1 cannot perform SOS mutagenesis, other species in the genus may have retained these genes and their ability to carry out SOS mutagenesis. We used ADP1 sequences to design Touchdown PCR amplification experiments and cloned *umuD* from *Acinetobacter* strains distributed throughout the

genus. Analyzing the *umuDC* genes from these strains showed that the *umuD* allele in ADP1 is common across the genus. Eight out of eleven strains analyzed had a *umuD* homolog possessing an extended 5' region. Three strains yielded no PCR product in Touchdown or inverse PCR experiments. The *umuD* sequences from these strains had high similarity (76-100% identity over 168 amino acids) to the *umuD* of ADP1. Additional data was collected from some strains on *ddrR* (an inducible SOS gene), the small *umuC* fragment, and the *umuD/ddrR* promoter region. These were also highly similar in the examined strains, differing only in several nucleotides. Analysis of the predicted UmuD gene products showed conservation in the UmuD self-cleavage catalytic residues Ser60/Lys97, the RecA-mediated self-cleavage site, and the recognition sites for Lon protease. Thus, we propose that the conservation of protein motifs suggest that *Acinetobacter* UmuD may self-cleave, although it is unclear what role UmuD might play after self-cleavage, given the absence of an intact *umuC* in at least 3 strains across this genus.

Accepted by: Janelle Hare, **Chair**

David K. Peyton

Acknowledgements

I would like to express my utmost gratitude to my advisor, Dr. Janelle Hare, for providing guidance and supporting me in this research. Her constantly valuable advice and assistance allowed me to not only build my knowledge and skills in the biological science field, but also to improve my ability in English. She was always patient with me. Without her help, I would not have accomplished my research.

I would also like to show my sincere gratitude to the committee members, Dr. Craig Tuerk and Dr. David Peyton, and to the BES Department graduate student coordinator, Dr. Geoff Gearner, for their valuable suggestions and comments that helped me make my thesis more complete. Also, I would like to thank my fellow students in the microbiological lab and the Morehead State University International Student Education Office for their continuous helps.

Finally, I am extremely grateful to my grandparents, parents and siblings for their unceasing love and encouragements. I am the most blessed person in the world to be their granddaughter, daughter, and sister. I would like to dedicate this thesis to them.

Table of Contents

List of Tables

Page

List of Figures

Introduction

DNA, which is the most fundamental component of life, can encode numerous functional protein products necessary for developing and maintaining life. However, DNA may suffer exogenous attacks under some conditions, causing alterations in DNA and interruptions in the encoding of functional proteins. In order to survive and replicate, a cell must possess and use a DNA damage sensing and repair system to respond to DNA damage. Numerous DNA repair systems are found in both eukaryotes and prokaryotes. This research examines the conservation and characterization of particular genes involved in one of the DNA repair systems in prokaryotes.

The introduction will discuss types of DNA damage sources and DNA response and repair systems. One of these DNA response systems is the error-prone repair system called SOS mutagenesis, which is well-studied in the gram-negative bacterium *Escherichia coli* (Friedberg *et al.*, 1995; Walker, 1984). The focus of my work was to examine the distribution and content of one of the critical operons involved in the SOS response in various *Acinetobacter* bacterial species, and how this operon compares to other bacterial operons responsible for error-prone SOS mutagenesis.

DNA (Deoxyribonucleic acid)

Deoxyribonucleic acid (DNA) is a long linear polymer, composes of four kinds of deoxynucleotides that carries genetic information to specify the amino acid sequences of proteins. All deoxynucleotides have a common structure composed of a phosphate group linked by a phosphodiester bond to a pentose (2'-deoxyribose) that in turn is linked to an organic base. The three-dimensional structure of DNA which consists of two long helical strands that are coiled around a common axis forming a double helix was first proposed by James D. Watson and Francis H. C. Crick in 1953 (Watson & Crick, 1953). Each strand of double-strand DNA is composed of four different types of monomers called deoxynucleotides. These four deoxynucleotides are adenine (A), thymine (T), cytosine (C), and guanine (G). Adenine and guanine are purines, which contain a pair of fused rings; cytosine and thymine are pyrimidines, which contain a single ring. Since native DNA is a double helix of complementary antiparallel chains, under normal conditions, an adenine pairs with a thymine and a guanine pairs with a cytosine; the pairs of purines and pyrimidines are held together by hydrogen bonds (Becherel & Fuchs, 1999).

Sources of DNA damage

Many artificial agents and natural factors in the environment can induce DNA damage (Ames, 1979; Hartl, 1999; Lodish *et al.*, 2001). Many chemical agents that can induce DNA mutations are described below (Singer & Kusmierek, 1982).

Base analogs can be incorporated into a DNA duplex in the course of normal replication. The most well-known base analog is 5'- bromouracil, which is an analog of thymine and can substitute for thymine in DNA. The keto group of 5'- bromouracil is complementary to adenine, so it can be incorporated into DNA by aligning to adenine during DNA replication; therefore the mismatch of adenine and 5'-bromouracil causes DNA mutation (Thinh & Griffiths, 1973).

Highly reactive chemicals, including nitrous acid and alkylating agents, can also damage DNA. Nitrous acid (HNO_2) can cause deamination by converting the amino group of a nucleotide into a keto group, such as a cytosine being converted to a uracil, which can then base pair with adenine and cause DNA mutation. Alkylating agents include ethylmethane sulfonate (EMS), nitrogen mustard and mitomycin C. EMS can add alkyl groups to DNA bases causing mismatches and leading to transition mutations. Nitrogen mustard, the cytotoxic chemotherapy agent for cancer treatment, can cause depurination via guanine alkylation. Mitomycin C can induce the formation of covalently cross-linked DNA, resulting in the DNA duplex distortion.

Other chemicals such as intercalating agents may cause base pair addition or deletion. The intercalating agent acridine is a 3-ringed molecule whose dimension is roughly the same as that of purine-pyrimidine pair. The intercalation causes addition or loss of a nucleotide during recombination or replication, thereby forming a frameshift. The most well-known intercalating agents are the nucleic acid stains ethidium bromide and acridine orange.

Besides the chemicals mentioned above, UV radiation and ionizing radiation absorbed by DNA are also responsible for DNA mutation. UV light irradiation can form linked pyrimidines (usually thymine-dimers) on adjacent strands. The chemical linkage brings the bases closer together, causing a distortion of the double helix structure and blocks transcription and DNA replication. This research focuses on the SOS operon that is induced by UV light damage in particular. Xeroderma pigmentosum, the human genetic disease which usually leads to severe sunburn or skin cancer at a young age, is the result of a deficiency in DNA polymerase η which is required for repairing UV-induced thymine dimers (Fujiwara *et al.*, 1999; Masutani *et al.*, 1999).

X-ray, alpha, beta particles and gamma ray are sources of ionizing radiation. Ionizing radiation can also result in several different types of mutations such as base substitution or single-strand or double-strand breaks, which cause the breakage of phosphodiester bonds.

Types of DNA damage

The main consequences of DNA damage caused by these extrinsic sources can be divided into either base substitution or base addition/deletion (Hartl, 1999; Lodish *et al.*, 2001; Maki, 2002). Base substitutions include transition mutations and transversion mutations. Transition mutations are base substitutions that either replace a purine with another purine or replace a pyrimidine with another pyrimidine, whereas transversion mutation occurs when a purine is replaced by a pyrimidine or a

pyrimidine is replaced by a purine. For damage by UV light, research shows that the targeted mutation frequencies are 42 % transitions and 58 % transversions at *cis*-syn thymine-dimer sites in cells (Banerjee *et al.*, 1988). Base substitutions can be the result of errors during DNA replication and chemical reactions such as oxidation.

Three types of outcomes are found in base substitution events. First of all is a silent mutation, which changes the nucleotide sequence without affecting the amino acid sequence, therefore, making no difference at the protein level. Second, a missense mutation changes both the nucleotide sequence and the resulting amino acid sequence. As the consequence of missense mutation, the encoded amino acid is different than that in the wild type, causing possible malfunction of the resulting protein. Sickle cell anemia is the well-known example of a disease caused by missense mutation. Lastly, nonsense mutation changes the nucleotide sequence, creating a new stop codon that disrupts the protein.

A frameshift mutation is an interruption in a reading frame when a single base or a number of bases not divisible by three is added to or deleted from the DNA sequence. Therefore, a frameshift mutation can shift the open reading frame (ORF) of the codons in the mRNA, rendering all of the amino acids downstream from the mutation site different from the original one. Frameshift mutations can be caused by intercalating mutagens such as ethidium bromide and acridine orange.

Lastly, some insertion mutations can also be caused by transposable elements (jumping genes). The insertion of transposable elements into the chromosome can cause either an interruption of the coding region, an intra-chromosomal duplication

(two copies of a jumping gene) or an unequal crossing-over. This mutation results in altered or abolished target gene function.

DNA sensing and repair systems

When DNA mutations occur, cells typically detect and respond to the damage via constitutive expression or induction of either an error-free repair system or an error-prone repair system to deal with the genotypic alteration.

Error-free repair systems

The error-free repair systems repair DNA sequences thoroughly to eliminate errors. Several repair systems involved in this system include proofreading, direct repair (direct reversal), and excision repair (Hartl, 1999; Lodish *et al.*, 2001).

The proofreading repair is a constitutive system. When a wrong base occasionally is inserted by DNA polymerase III during DNA synthesis, the 3' to 5' exonuclease activity of DNA polymerase I can remove a mismatched base at the 3' growing end of a synthetic primer-template complex. Because of the "proofreading" function provided by DNA polymerase I, this is system is called the proofreading repair system.

The direct repair system, or the photoreactivation repair system, acts to reverse the mutagenic event. This system applies several cellular enzymes such as photolyase, DNA glycosylase, AP endonuclease, and methyltransferase to repair UV-induced DNA damage. The photolyase can recognize pyrimidine dimers, bind to the

photodimer and break the bonds that join the pyrimidines in the dimer in the presence of visible light. The alkylated or deaminated bases can be removed from DNA by special DNA glycosylases, leaving an apurinic or apyrimidinic site in the DNA. These sites can be then recognized by AP endonucleases which in turn remove the deoxyribose-phosphate from the backbone. The gap caused by the process can be then repaired by DNA polymerase I and DNA ligase. Additionally, the methyl transferases can remove the extra methyl groups from nucleotides.

The excision repair system can be classified into three subtypes: base-excision repair, nucleotide-excision repair and mismatch repair. Cells can use excision repair to fix DNA regions containing chemically modified bases (chemical adducts) that distort the normal shape of DNA locally. Proteins that are involved in this repair system can slide along the surface of double-stranded DNA looking for bulges. Once they encounter a lesion, they will bind to the damaged strand and carry out the removal of a short stretch of nucleotides containing a major distortion in the DNA double helix. After this, DNA polymerase I and DNA ligase will seal the remaining nick.

Base-excision repair can repair a single mismatched base. DNA glycosylase and AP endonucleases play important roles in the base-excision repair system. The nucleotide-excision repair can repair the damage that affects oligonucleotides (2 to 30 bases) such as thymine dimers. The difference between these two repair systems is that the nucleotide-excision repair removes oligonucleotides rather than just a single base.

The mismatch repair system can repair either base substitution mutations, which usually start as a single mismatched base in double-stranded DNA. These mismatched bases are often detected and corrected by a mismatch repair system through excision of a single-stranded region containing the mismatched base and re-synthesis via using the remaining strand as the template. The products of three genes, *mutH, mutL* and *mutS,* participate in the mismatch repair system (Cox, 1997). The proteins MutH and MutS can detect the mismatched base and bring double-stranded DNA together, forming a DNA loop via facilitation from MutL (Feng *et al.*, 1996). The enzymatically active MutH nicks on one strand containing the mismatched base, which is then excised. After excision, DNA polymerase fills in the gap using the remaining strand as a template, thus eliminating the mispair.

The mutation frequency of one thymine dimer in cells depends on the type of repair systems that cells used. The range of overall mutation frequency of one thymine dimer is 7~11 % after UV-induced cells apply SOS mutagenesis repair, while the overall mutation frequency is only 4 % in non UV-induced cells and the absence of SOS induction (Banerjee *et al.*, 1990).

Error-prone repair system

When cells are suffering lethal, massive DNA damage and error-free repair systems cannot repair the damage, cells may apply another repair system, an error-prone repair system, to overcome DNA damage. In other words, this "error-prone"

system permits the cells to tolerate the DNA damage via bypassing the lesion site with a lower fidelity of repair, rather than actually repair it accurately.

The SOS response was first discovered by and named by Miroslav Radman in 1974 (Radman, 1974). When DNA damage is so extensive that replication may occur before constitutive mechanisms (error-free mechanisms) can repair all the damage, an *umuDC*-dependent alternative process called translesion DNA synthesis (TLS) is induced. TLS is a mechanism whereby cells pay the cost of an elevated mutation rate in exchange for increased survival (Friedberg *et al.*, 1995; Sutton *et al.*, 2000). With regard to the mutations that result, this error-prone inducible system is also called "SOS mutagenesis" (Becherel & Fuchs, 1999). The SOS system can help avoid a lethal interruption of DNA replication resulting from the inability of the DNA polymerase III to replicate through lesions such as abasic sites (Tang *et al.*, 1998, 1999; Reuven *et al.*, 1998, 1999), UV-induced thymine-dimer and photoproducts (Tang *et al.*, 2000) in the template DNA. Moreover, it can be regulated at both transcriptional and post-translational levels, involving the repressing and activating of SOS genes and the subsequent modification of encoded SOS proteins.

The SOS mutagenesis system

SOS mutagenesis in Escherichia coli

Our understanding of the SOS mutagenesis system has been established by intensive investigation in *Escherichia coli* (Friedberg *et al.*, 1995; Khil & Camerini-Otero, 2002; Little & Mount, 1982; Sutton *et al.*, 2000; Walker, 1985).

Even though there are over 1000 genes involved in the overall DNA damage response of *E. coli* (Khil & Camerini-Otero, 2002), the SOS mutagenesis system only requires about 30 unlinked genes (e.g., *recA, lexA, umuDC, polB, recN, sulA, uvrAB and uvrD*) (Fernandez *et al.*, 2000; Friedberg & Walker, 1995) that compose a SOS regulon which responds to DNA damage. Although every encoded protein from these SOS genes has its own specific function, a few genes are especially important in SOS mutagenesis. The proteins LexA and RecA are in charge of the inhibition and activation, respectively, of SOS genes such as the *umuC* and *umuD* genes (Walker, 1996). The *umuC* and *umuD* genes that encode the UmuD and UmuC subunits of DNA polymerase V are the most critical effectors in this repair system (McNally *et al.*, 1990; Sommer *et al.*, 1993). DNA polymerase V can replicate past the lesions in DNA strands (TLS) without a proofreading function (Maor-Shoshai *et al.*, 2000; Tang *et al.*, 1998, 2000) and thus causes SOS mutagenesis (Sutton *et al.*, 1999). This transient increase in the mutation frequency of chromosomal genes following induction of the SOS mutagenesis is characterized by a striking increase in transversions (Fijalkowska *et al.*, 1997).

Upstream of the coding region of SOS genes in *E. coli*, there is a promoter-operator complex having the consensus sequence CTGTATATAAAAACAG, which is referred to as a SOS box (Mount *et al.*, 1972). The locations of SOS boxes vary with respect to transcriptional start sites; they may overlap either with the -35 or -10 promoter consensus elements (the -10 for *lexA* and *umuDC*), be between the -35 and -10 elements, or be downstream of the -10 element (Walker, 1996). LexA binds the

SOS box in the promoters of various SOS genes and negatively regulates their transcription when DNA damage is absent (Mount *et al.*, 1972). The transcription of the *umuDC* operon begins at an adenine residue in the SOS box of *E. coli*. An SOS box in the promoters of SOS genes is conserved in distantly related bacteria such as gram-negative bacteria, gram-positive bacteria and cyanobacteria, although the nucleotide sequences may differ among these bacteria.

The expression of many genes in the SOS regulon is controlled by the LexA protein, a transcriptional repressor. The functional LexA protein is a dimer (Schnarr *et al.*, 1985; Thliveris *et al.*, 1991). Each monomer of LexA contains a DNA binding domain at its amino-terminal and dimerization domain at its carboxyl-terminal; unless the dimer is formed, the DNA binding domain will not bind to the SOS box in the promoters of SOS genes. Under normal conditions, the LexA protein binds to the *umuDC* promoter and repressively regulates its transcription by interfering with the binding of RNA polymerase (Brent *et al.*, 1981; Little *et al.*, 1981; Kitagawa *et al.*, 1985).

Another regulatory protein involved in the SOS response is RecA. The general function of RecA in the cell is to carry out the repair of daughter-strand gaps, double-strand breaks (Cox, 1991), and to conduct the SOS response (Friedberg & Walker, 1995). With regard to the SOS mutagenesis system, RecA is necessary for facilitating both the intramolecular autodigestion of LexA (Story *et al.*, 1993) and the intermolecular UmuD protein self-cleavage (Battista *et al.*, 1990). The SOS response begins after DNA damage occurs, when single-stranded DNA is formed (Higashitani

et al., 1992), and normal DNA replication is blocked (Craig & Roberts, 1981; Sassanfar & Roberts, 1990). At this time, RecA binds to the single-stranded DNA, forming a helical, multimeric nucleoprotein filament (Craig & Roberts, 1981; Sassanfar & Roberts, 1990) and achieving an activated state, called RecA*. RecA*, a co-protease, can facilitate the proteolytic cleavage of the bond between alanine-84 and glycine-85 of LexA (Little *et al.*, 1980). As the result of LexA autodigestion (Little, 1984), the degraded LexA repressor is removed from the SOS box and no longer represses the transcription of SOS genes such as *umuDC*.

The term Umu is an abbreviation of "UV mutagenesis." The UmuD and UmuC proteins are the products of *umuDC* genes and play a critical role in the SOS mutagenesis system. Without the repression of LexA in their promoter region, the proteins UmuD and UmuC are made, forming a UmuD homodimer ($UmuD_2$) and UmuC. The $UmuD_2$ homodimer acts as a checkpoint inhibitor of cell division, allowing additional time for cells to repair accurately by using error-free repair system such as excision repair, before post-translational regulation is induced (Opperman *et al.*, 1999). With the facilitation of RecA*, $UmuD_2$ self-cleaves and removes its 24 amino-terminal residues to yield a functional $UmuD'_2$ protein in ~30 minutes (Burckhardt *et al.*, 1988; Nohmi *et al.*, 1988; Shinagawa *et al.*, 1988). The functional $UmuD'_2$ homodimer interacts with UmuC to form a functional SOS protein: the $UmuD'_2C$ complex (Woodgate *et al.*, 1989; Reuven *et al.*, 1998, 1999; Tang *et al.*, 1998, 1999, 2000), otherwise known as DNA polymerase V (Figure 1).

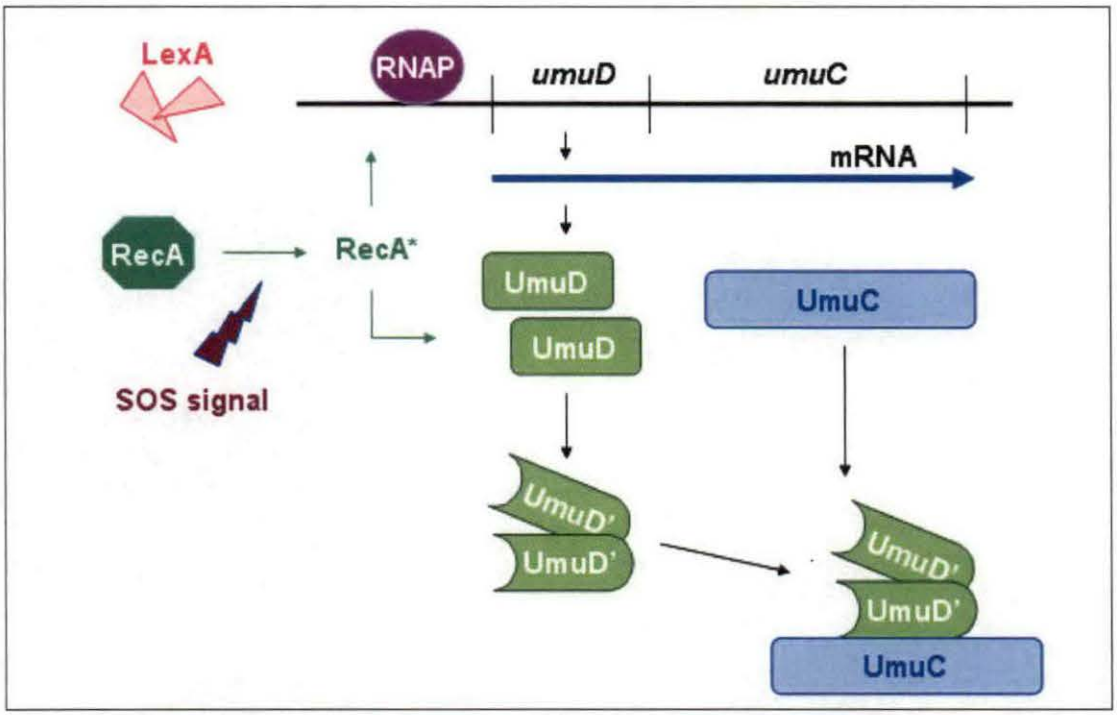


Figure 1. The induction and formation of DNA polymerase V in *E. coli* after DNA damage.

DNA polymerase V is a member of the Y-family of DNA polymerases involved in translesion DNA synthesis. The biochemical properties of these TLS polymerases appear to be conserved from prokaryotic DNA polymerase V to eukaryotic DNA polymerases η, ι, κ, and Rev1 (Hübscher *et al.*, 2002).

DNA polymerase V plays a central role to help cells survive under the deleterious conditions of DNA damage such as the UV-induced T-T dimer. When the replication forks encounter the lesions, DNA polymerase III will be blocked, stalling the replication process. After the SOS response, the $UmuD'_2C$ complex can force DNA replication through the lesion, however with a cost of making errors, because it has a low replication fidelity and therefore adds wrong nucleotides such as G-T opposite a T-T dimer.

In *E. coli*, the intact UmuD protein is 140 amino acids and shares carboxyl-terminal homology with LexA (Walker, 1996). The UmuC protein is 423 amino acids. The post-translational modification of UmuD and UmuC proteins contains multiple steps. In order to form a functional SOS protein complex, the UmuD protein has to self-cleave. McDonald generated compatible plasmids that expressed a UmuD protein with mutations at either the active site or the cleavage site of the protein. Neither mutant protein underwent cleavage when individually expressed in the same cell, but when they were co-expressed, the UmuD′ protein was detected. This indicates that intermolecular cleavage occurred (McDonald *et al.*, 1998).

Additionally, UmuD self-cleavage needs help from RecA* both *in vivo* (Shinagawa *et al.*, 1988) and *in vitro* (Battista *et al.*, 1990; Burckhardt *et al.*, 1988).

According to an NMR study of UmuD, the catalytically active serine-60 and lysine-97 residues are far apart (Lee *et al.*, 1994). The action of RecA* seems to function by bringing these two residues closer together, so catalysis can be initiated between the active site and cleavage site of the UmuD protein under physiological conditions (Sutton *et al.*, 2001).

The β-clamp/γ-clamp loader of DNA polymerase III also interacts with DNA polymerase V, RecA and single-stranded-binding protein, forming a "mutasome" (Rajagopalan *et al.*, 1992; Tang *et al.*, 1998, 1999) at the DNA damage site in order to provide an optimal condition for translesion replication (Goodman & Woodgate, 2000).

Other proteins are involved in the negative regulation of the error-prone DNA repair system. Because the low-replication-fidelity of DNA polymerase V increases mutation rate, the formation of DNA polymerase V is the last resort for cells. In order to reduce and control the intracellular levels of UmuD and UmuD', two proteolytic proteins, Lon and ClpXP, regulate and degrade UmuD and UmuD' by recognizing certain recognition sites.

Lon, an ATP-dependent serine protease, recognizes two sites (primary and secondary sites) within the amino-terminal end of UmuD (Figure 2), as demonstrated by site-directed mutagenesis. The primary and secondary Lon-recognition sites are Phe15-Pro16-Leu17-Phe18 and Phe26-Pro27-Ser28-Pro29, respectively, and both sites are required for the efficient degradation of UmuD (Gonzalez *et al.*, 1998). Regulating the cellular level of UmuD via Lon-mediated proteolysis provides active

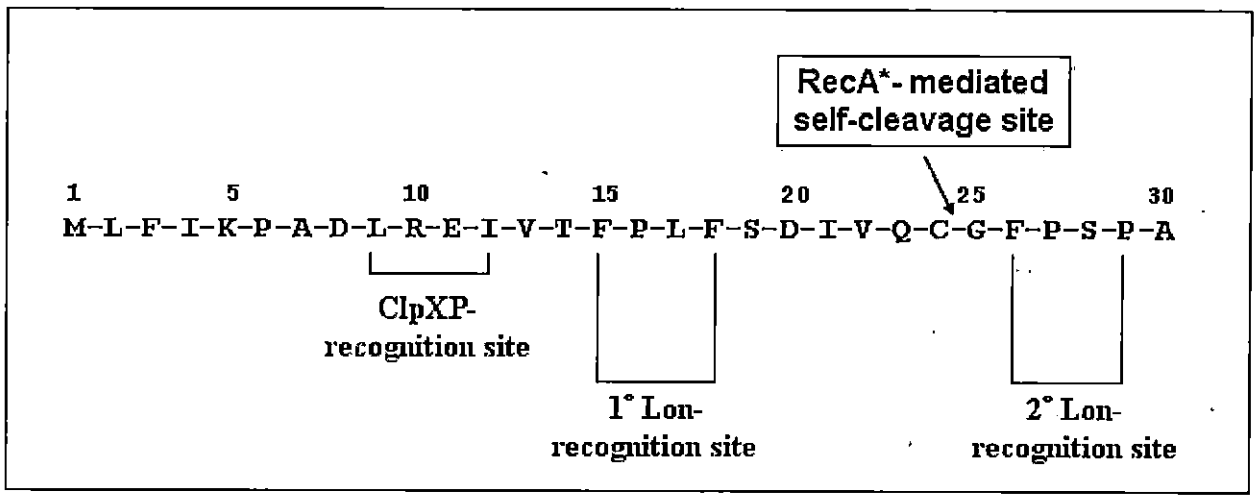


Figure 2. Functional motifs within the N-terminus of UmuD in *E. coli.* Here shows the recognition sites for ClpXP and Lon proteases, and the RecA* self-cleavage site. This figure is adapted from Gonzalez, *et. al.*, 2002.

and more stable UmuD′ proteins. However, UmuD and UmuD′ preferentially interact to form UmuD/UmuD′ heterodimers rather than UmuD′$_2$ homodimers and these heterodimers are functionally inactive. This is because ClpXP, a heterooligomeric ATP-dependent serine protease composed of a proteolytic subunit (ClpP) and a protease ATP-binding subunit (ClpX), can direct the rapid degradation of UmuD′ in a UmuD/UmuD′ heterodimer (Gonzalez *et al.*, 2000). However, the ClpXP recognizes the amino-terminal 24 amino acids on the UmuD protomer, which is Leu9-Arg10-Glu11-Ile12 (Figure 2), and thus directs its action solely at the UmuD′ subunit of the heterodimer rather than UmuD subunit (Battista *et al.*, 1990; Gonzalez *et al.*, 2000).

To summarize, Lon and ClpXP are negative regulators for SOS mutagenesis. ClpXP can keep the basal levels of UmuD′ at a minimum in undamaged cells and also act in damaged cells to reduce the elevated levels of mutagenically active UmuD′ protein (Gonzalez *et al.*, 1998). The proteolytic mechanism used by both Lon and ClpXP allows cells to reduce their intracellular levels of the mutagenetically active UmuD and UmuD′ proteins and thereby return to a resting state once error-prone DNA repair has occurred (Peat *et al.*, 1996).

A recently discovered SOS protein, DinI, can also modulate the coprotease activity of the RecA protein (Yasuda *et al.*, 1998). DinI inhibits RecA-mediated autodigestion of UmuD to UmuD′. The induced SOS mutagenesis frequency in a *dinI*$^-$ strain is thus higher than in a *dinI*$^+$ strain because more UmuD becomes UmuD′.

SOS mutagenesis systems in other bacterial model systems

Besides being well-studied in *E. coli*, SOS mutagenesis is also studied in many different bacteria. In some distantly related bacteria species, such as *Bacillus* (Cheo *et al.*, 1991; Davis *et al.*, 2002), *Mycobacterium* (Movahedzadeh *et al.*, 1997; Brooks *et al.*, 2001) and *Xanthomonas* (Yang *et al.*, 2002; Yang *et al.*, 2005), the SOS mutagenesis systems generally resemble the SOS system present in *E. coli.* For example, the SOS box and LexA proteins are functionally and generally conserved in these bacteria.

The presence of the SOS box is required for the negative regulation of SOS gene transcription via the promoter regions of various SOS genes in *E. coli.* However, the specific requirement of LexA and the SOS box in other bacteria are variable in some ways. The number of *lexA* genes present in bacteria is variable, ranging from 0 to 2 (Campoy *et al.*, 2002; Jara *et al.*, 2003). Moreover, the conserved SOS box sequence also exists in many bacteria, even though the specific sequences of the SOS box present between different bacterial classes vary, such as TTAG(N_6)TACTA for *Xylella fastidiosa* (Campoy *et al.*, 2002), or CGAACRNRYGTTCYC for *Bacillus subtilis* (Winterling *et al.*, 1998).

Additionally, research has also revealed a variety of differences in the way that bacteria respond to DNA damage, including *Acinetobacter baylyi* strain ADP1. In order to increase our understanding about SOS mutagenesis, it is worthwhile to study SOS mutagenesis-related genes overall in other model systems such as *Acinetobacter baylyi* strain ADP1.

***Acinetobacter* biology**

Members of the non-motile, coccobacillary, strictly aerobic gram-negative bacterial genus *Acinetobacter* are abundant in most environments, including water, soil, living organisms and human skin (Juni, 1978). Most *Acinetobacter* species are non-pathogenic, except for *A. baumannii, A. ursingii, and A. schindlerii* (Nemec *et al.*, 2000). These are opportunistic pathogens for immunocompromised people and are responsible for nosocomial infections. The genus *Acinetobacter* is classified in Phylum *Proteobacteria*, Class *Gammaproteobacteria*, Order *Pseudomonadales* and Family *Moraxellaceae*. *Acinetobacter* contains 32 genomic species (Nemec *et al.*, 2000, 2001, 2003), 17 species of which were named by 2003 (Carr *et al.*, 2003). Even though *Acinetobacter* and *Escherichia* are in the same Class, they differ in their Orders. *Escherichia* is classified in Order *Enterobacteriaceae*. Therefore, these two species are not closely related.

The *Acinetobacter* species strain ADP1, recently defined as *Acinetobacter baylyi* strain ADP1 (Vaneechoutte *et al.*, 2006), has been sequenced and possesses an average bacterial genome size of 3.7 Mb (Barbe *et al.*, 2004). Its GC content is 40.3%, while the GC content range of all the *Acinetobacter* species is 38-47% (Barbe *et al.*, 2004). *Acinetobacter baylyi* strain ADP1 is highly competent for natural genetic transformation (Carr *et al.*, 2003; Barbe *et al.*, 2004) and this phenotype is proposed to be the trait which best differentiates this species from all other *Acinetobacter* species (Vaneechoutte *et al.*, 2006). Moreover, because of this ability, the characteristics of easy, rapid lab growth and possession of no obvious pathogenicity

or virulence factors, ADP1 is considered to be an ideal organism model for genetic analysis and genome engineering, similar to *E. coli* (Metzgar & Bacher, 2004).

DNA damage response genes and characteristics found in ADP1

Similar features of DNA damage response in ADP1

At least two genes in ADP1 are induced by DNA damage: *ddrR* and *recA*. In ADP1, a novel genetic locus with no homologs in other bacteria is located upstream of the *umuDC* operon (Whitworth & Gregg-Jolly, 2000). Because this genetic locus is induced (~10-fold) in response to DNA damage, it was named *ddrR* (DNA damage response gene R) (Figure 3) (Hare *et al.*, 2006). This DNA damage induction is dependent on *recA* (Whitworth & Gregg-Jolly, 2000). This is consistent with the model in *E. coli.* However, UmuDC may trancriptionally regulate *ddr*R, because the insertion of a *lacZ*:Km^R resistance cassette into *umuD* reduces *ddrR* induction by 83% (Hare *et al.*, 2006). This activity of UmuD is unprecendented.

Also as in the *E. coli* model, *recA* is induced by DNA damage in ADP1 (Rauch *et al.*, 1996). However, unlike *ddrR*, the induction of *recA* is not dependent on a functional RecA protein (Rauch *et al.*, 1996). In other words, *recA* transcription is induced in response to DNA damage in ADP1, although the induction does not require the RecA protein. Therefore, *recA* expression in ADP1 appears to be regulated in a novel and unknown mechanism, possibly involving a non-LexA-like repressor.

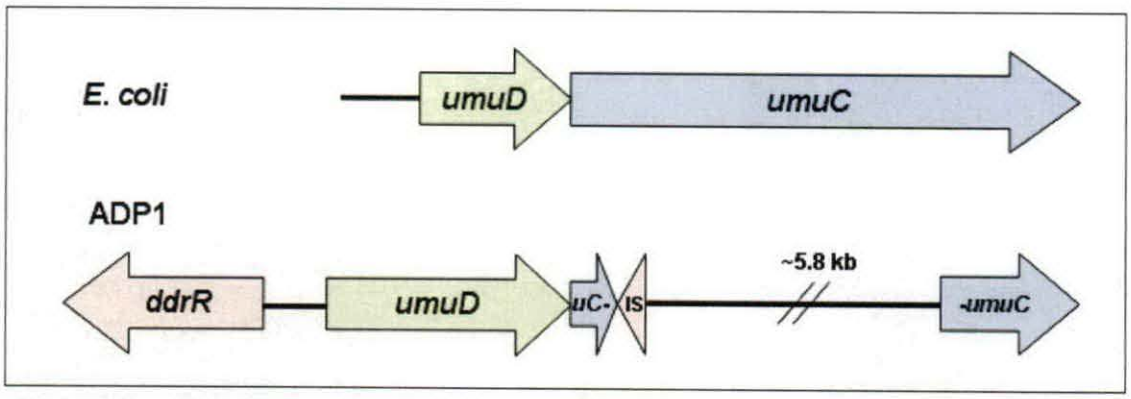


Figure 3. The comparison of the *umuDC* region in *Escherichia coli* and in *Acinetobacter baylyi* strain ADP1. The *umuD* homolog is 1.5-fold longer and *umuC* homolog is mutated in *Acinetobacter* ADP1. The mutated *umuC* homolog is separated into two fragments; the small one is right next to 3' region of *umuD* homolog, the large one is ~5.8kb downstream of *umuD* and the middle region has no homologous sequence found in ADP1. A novel genetic locus, *ddrR* (DNA-damage inducible gene R), is found in ADP1 without homologs in other bacteria. This figure is adapted from Hare, 2006.

umuD induction is independent of functional RecA and DNA damage

In *E. coli*, the *umuDC* operon is induced about 15-30 fold in response to DNA damage (Friedberg & Walker, 1995). However, the *umuD* gene is not regulated by either DNA damage or functional RecA in ADP1 (Hare *et al.*, 2006). Furthermore, in ADP1, the analysis of a *umuD::lacZ* fusion reveals that *umuD* is expressed at high levels in both the absence and presence of DNA damage. This may be because no LexA is found in ADP1 (Hare *et al.*, 2006). However, as functional DNA polymerase V requires both UmuD and UmuC, no DNA polymerase V is encoded in ADP1, because *umuC* is truncated.

Absence of both LexA and the SOS box

Interestingly, however, there is no SOS box or shared nucleotide sequence found in the *umuDC* promoter region (Hare *et al.*, 2006) or in the *recA* promoter region in ADP1 (Rauch *et al.*, 1996). Moreover, no LexA homolog exists in ADP1. A BLAST of the ADP1 genome with the *E. coli lexA* gene identifies *umuD* as the closest match (37 % a.a. identity).

The "extra-long" UmuD is found in ADP1 and in other microbes

The *umuD* genes in most bacterial species resemble those in *E. coli*, encoding the normal size UmuD proteins (~140 amino acids) in enterics and non-enterics alike, e.g. *Citrobacter, Vibrio cholerae, Providencia rettgeri, and Prochlorococcus marinus.*

The *umuD* gene in *Acinetobacter* sp. strain ADP1 is 612 bp (referred to in Barbe' (2004) as ACAID2729; chromosome coordinates 2673896 to 2674507) in length which encodes 203 amino acids (Figure 3). Therefore, the *umuD* open reading

frame of ADP1 is 1.5-fold longer than the *E. coli umuD* because an additional 58 amino acids are observed at the amino-terminal UmuD protein. Moreover, the self-cleavage site, cysteine-24/glysine-25 (Perry *et al.*, 1985), is replaced by alanine-24/glycine-25, although other cleavable UmuD proteins in other bacteria also possess AG also at this location. Two amino acids, leucine-101 and arginine-102, which are required for efficient UmuD self-cleavage in *E. coli* (Sutton *et al.*, 2001) are replaced by isoleucine-163 and aspartate-164 in ADP1. An extra five amino acids adjacent to aspartate-164 are uniquely found in ADP1 and not the other "extra-long" UmuD proteins (Hare *et al.*, 2006). However, the catalytic residues serine-60 and lysine-97 (Nohmi *et al.*, 1988), which are required for UmuD self-cleavage in *E. coli*, are both conserved in ADP1.

However, some bacterial species besides ADP1 have a bigger UmuD than the *E. coli* UmuD (Hare *et al.*, 2006), including *Synechococcus elongatus, Chromobacterium violaceum, Legionella pneumophila, and Thiomicrospira crunogena.* The "extra-long" UmuD homologs in these bacteria have not been studied, but were identified in genomes that were sequenced (Hare *et al.*, 2006). *Synechococcus elongatus* is a unicellular freshwater cyanobacterium which is transformable by exogenously added DNA. Many genetic tools have been developed for it in the past several decades. The extra-long UmuD in *Synechococcus elongatus* is 186 amino acids in length (Sugita *et al.*, 2007). *Chromobacterium violaceum*, a gram-negative ubiquitous bacterium that is abundantly found in subtropical and tropical ecosystems, also has an extra-long UmuD which is 198 amino acids in length

(Vasconcelos *et al.*, 2003). The extra-long UmuD protein encoded in the gram-negative, but opportunistic bacterium, *Legionella pneumophila*, which is best known for causing Legionnaire's disease, is 168 amino acids (Chien *et al.*, 2004). *Thiomicrospira crunogena* is also a ubiquitous bacterium first isolated from East Pacific Rise (Jannasch & Mottl, 1985). The UmuD in *Thiomicrospira crunogena* is 206 amino acids in length (Scott *et al.*, 2006). One interesting thing is that, although these bacteria all share a very unusual UmuD protein, they belong to diverse taxa. One possible reason may be that the ancestors of these bacteria including *E. coli* possessed a bigger UmuD rather than a smaller one; under the evolutionary pressure, however, some of them were mutated by unknown reasons.

umuC is mutated

In contrast to *umuD*, the *umuC* gene in ADP1 is present in two fragments which together contain only 43% of the length of the *umuC* gene in *E. coli* (Hare *et al.*, 2006) (Figure 3). The truncated *umuC* gene fragment only encodes thirty-nine amino acids homologous to the amino-terminal end of UmuC (with 85% identity) and is interrupted by a fragment of a putative transposase gene found in the opposite orientation. This putative transposase gene is only 72 bp long and exhibits homology to IS*Ehe*3 (Hare *et al.*, 2006). Furthermore, a 348-bp fragment of *umuC* is found at 5.8 kbp downstream of *umuDC,* with only 28% identity to UmuC in its 114 encoded amino acids. A gene with homology to the transcriptional regulator *lysR* is also found in between these two fragments.

Conserved characteristics found in ADP1

Even though there are many unusual features regarding the response to DNA damage found in the ADP1 model system, some characteristics are still conserved in ADP1. Homologs of the ATP-dependent serine proteases Lon and ClpXP are also found in *Acinetobacter* species. In *E. coli,* Lon and ClpXP can recognize certain residues in the amino-terminus of UmuD, which are close to the RecA-mediated self-cleavage site, cysteine-24/glycine-25. Therefore, it will be interesting to learn if the "extra-long" UmuD in ADP1 or other *Acinetobacter* species also possesses these recognition sites which are close to its putative RecA-mediated self-cleavage site, alanine-24/glycine-25.

Goals for this project

As already mentioned above, some unique features in response to DNA damage have been revealed by studying in *Acinetobacter baylyi* strain ADP1. However, not much information is known about the DNA damage response systems and their related genes in all other *Acinetobacter* species. Do any of the unique features of ADP1's response to DNA damage come from its unusual *umuDC* operon structure and content? My goal in this project was to investigate the *umuDC* operon within the *Acinetobacter* genus.

There are thirty-two different genomic groups in *Acinetobacter* genus (Carr *et al.*, 2003); however, whether they possess the same unusual *umuDC* alleles as found in ADP1 is unknown. In this project, I would like to examine these aspects of the

umuDC operon: (1) Do the *umuD* and *umuC* genes from other *Acinetobacter* strains also have an elongated *umuD* and a mutated, split *umuC* as those present in ADP1? The result may demonstrate whether the unusual *umuD* and *umuC* genes are conserved within the *Acinetobacter* genus or unique to ADP1. (2) Do the *umuD* and *umuC* genes from *Acinetobacter* strains other than ADP1 possess the intact *umuDC* operons, similar to *E. coli*? In *E. coli*, the *umuD* and *umuC* genes are next to each other, so they are transcribed from the same promoter. Thus, if the *umuD* and *umuC* genes from other *Acinetobacter* strains are next to each other, we can say they are likely regulated coordinately. (3) Do all of the examined strains with *umuD* also have *umuC*? Both *umuC* and *umuD* are present in *E. coli*; therefore, the translational protein product of the *umuDC* operon can form a functional DNA polymerase V that executes error-prone DNA repair. In ADP1, however, there is no intact *umuC* that can encode the functional UmuC protein.

Materials & Methods

Acinetobacter sp. strains

All the examined strains used in this project, with the exception of ADP1, were purchased from the American Type Culture Collection (ATCC) to avoid analyzing any mutations arising through long-term laboratory maintenance. These strains are stored in a -80°C freezer.

The examined strains used in this project are listed in Table 1: *Acinetobacter conjunctivae* strain TU14 (ATCC 17905), *Acinetobacter haemolyticus* (ATCC 17906), *Acinetobacter junii* (ATCC 17908), *Acinetobacter johnsonii* (ATCC 17909), *Acinetobacter calcoaceticus* (ATCC 23055), *Acinetobacter radioresistens* (ATCC 43998), *Acinetobacter anitratus* (ATCC 49137), *Acinetobacter* genomospecies 3 (ATCC 19004), *Acinetobacter* genomospecies 6 (ATCC 17979), *Acinetobacter* genomospecies 9 (ATCC 9957), *Acinetobacter* genomospecies 10 (ATCC 17924), *Acinetobacter* genomospecies 14 (ATCC 51552), *Acinetobacter* genomospecies 16 (ATCC17988), and *Acinetobacter baylyi* (ADP1).

Growth conditions

All of the examined strains grow well on either nutrient agar (nutrient broth) or tryptic soy agar (TSA) or tryptic soy broth (TSB), including *Acinetobacter* sp. genomospecies 6, although its suggested medium is brain heart infusion (BHI) agar

Table 1. The optimal growth conditions for *Acinetobacter* strains suggested by ATCC.

Strain (ATCC#)	Medium	Temperature (°C)
Acinetobacter sp. genomospecies 3 (19004)	NB	37.0
Acinetobacter sp. genomospecies 6 (17979)	BHI	37.0
Acinetobacter sp. genomospecies 9 (9957)	NB	37.0
Acinetobacter sp. genomospecies 10 (17924)	NB	26.0
Acinetobacter sp. genomospecies 14 (51552)	NB	30.0
Acinetobacter sp. genomospecies 16 (17988)	NB	26.0
Acinetobacter anitratus (49137)	NB	30.0
Acinetobacter calcoaceticus (23055)	NB	30.0
Acinetobacter conjunctivae TU14 (17905)	NB	26.0
Acinetobacter haemolyticus (17906)	NB	26.0
Acinetobacter johnsonii (17909)	NB	26.0
Acinetobacter junii (17908)	NB	26.0
Acinetobacter radioresistens (43998)	NB	30.0

NB = Nutrient Broth

BHI = Brain Heart Infusion

or BHI broth. However, the growth temperatures for these strains vary from 26 to 37°C (Table1).

Extraction of genomic DNA

The cells were grown overnight (16-20 hours), shaking in 5 ml nutrient broth media (or TSB) at 180 rpm at temperatures ranging from 26°C to 37°C (depending on the required growth temperature of each strain). On the next day, the overnight cultures were pelleted at 14,000g for 5 minutes. Genomic DNA purifications were performed using the Wizard® Genomic DNA Purification Kit (Promega). After DNA purifications, the genomic DNA of each examined strain was incubated at 65°C for 15 minutes and 3µl was run on an 0.8% agarose gel in order to determine whether the DNA purifications yielded high quality DNA. Also, control PCR amplification was performed with a set of primers, 16S rRNA For and 16S rRNA Rev, to confirm that PCR amplification can be performed on this genomic DNA. The size of this product is 575 bp in *Acinetobacter* species. The purified genomic DNA preparations were maintained at 4°C for future use.

PCR amplification

PCR Master Mix, 2X (Promega) was used as the PCR reagent for all of PCR amplifications. It is a premixed, ready-to-use solution containing *Taq* DNA polymerase, dNTPs, $MgCl_2$ and reaction buffers at optimal concentrations for

efficient amplification of DNA templates by PCR. The amount of DNA template for PCR amplification was 0.5-1.0 µl in a 25-50 µl reaction.

According to the protocol of PCR Master Mix, 2X, the PCR amplifications were held at 95°C for 2 minutes, then the DNA templates were denatured at 95°C for 30 seconds (Table 2), annealed at the 5°C below the calculated T_m (°C) for 30 seconds, and extended at 72°C for the most effective time depending on the length of the target region. In the extension step, 1 kb of expected DNA product requires an extension time of 1 minute. For each run of PCR amplification, 30 cycles were always run and ended with a final extension at 72°C for 5 minutes. The PCR products were always stored at 4°C before running on gels.

A positive control (containing ADP1 genomic DNA as the template) and a negative control (nuclease-free water as template) were included in each PCR amplification experiment.

Touchdown PCR amplification

PCR Master Mix, 2X (Promega) was used as the PCR reagent for all of PCR amplifications. The amount of DNA template for PCR amplification was 0.5-1.0 µl in a 25-50 µl reaction.

According to the protocol of PCR Master Mix, 2X and the recommendation from the Touchdown PCR procedure (Don *et al.*, 1991), the Touchdown PCR amplifications were held at 95°C for 2 minutes and the examined DNA templates were then denatured at 95°C for 30 seconds. The difference between regular PCR and

Table 2. PCR primers designed upon ADP1 and *E. coli* genomic DNA sequences. Primers were designed from the regions of (1) *umuD* and *umuC* in both ADP1 and *E. coli*; (2) the *ddrR* genetic locus in ADP1; (3) a *lysR*-like gene in ADP1; and (4) large fragment of *umuC* in ADP1.

Primer	Sequence	Base pairs	T_m (°C)	Binding site
CL0	5'-AGCCAACTAAAGTCATTCG-3'	19	54	*umuD* in ADP1
CL1	5'-GATCTCCGCCGAAATTGTG-3'	18	54	*umuD* in *E.coli*
CL2	5'-GCCTATGTCAGTTGTGAG-3'	18	54	*umuC* in ADP1
CL3	5'-GTTGTATAACGTGGTGAAAGC-3'	21	60	*umuD* in *E.coli*
CL4	5'-CCTGCTTATGCAATGACAG-3'	19	56	*lysR*-like in ADP1
CL5	5'-AGATCACGAGTTCTTGACC-3'	19	56	*umuD* in ADP1
CL6.2	5'-TTATGCCAGCTGTGAGACG-3'	19	54	*umuC* in *E.coli*
CL7	5'-ATCGCCTTAACGACGTGG-3'	18	56	*umuD* in ADP1
CL8	5'-CAGAGTCTGGATCTGAATG-3'	19	56	middle of *umuD* in ADP1
CL9.2	5'-ATCACGATATCACCTGC-3'	17	50	*umuD* in ADP1
CL10.2	5'-TGCGATGGTTGACAATGAG-3'	19	56	*umuD* in ADP1
A	5'-TAAGCATGTAGCTCTTGGG-3'	19	56	*ddr*R in ADP1
B	5'-CTTGAAAGTACAATCACAG-3'	19	52	*umuC* in ADP1
K	5'-TAACGCATAGGTTTCAGATTG-3'	21	58	*umuD* in ADP1
L	5'-AGTCATGAGTCAGAG-3'	15	44	*umuC* in ADP1
M	5'-TTGTCATCGAATAAATTGAGC-3'	21	56	middle of *umuC* in *E.coli*
N	5'-ATAGTGTTGGTATGATGCG-3'	19	54	middle of *lysR*-like in ADP1
O	5'-AATTTACACGTAAAAGAGC-3'	19	50	large piece *umuC* in ADP1
P	5'-TTGAAATTAATCAGGCG-3'	17	46	large piece *umuC* in ADP1
16S rRNA For	5'-GCACCTGTATGTAGATTCC-3'	19	56	16S rRNA of ADP1
16S rRNA Rev	5'-TACTCGCAGAATAAGCACC-3'	19	56	16S rRNA of ADP1

Touchdown PCR is in the annealing step. For the Touchdown PCR amplification, the annealing step was started at 3°C below the optimal Tm (°C) for 30 seconds and the annealing temperature was decreased by 1°C every 3 cycles (the range of the highest temperature to the lowest temperature was 11°C), and eventually run 10 cycles at the lowest temperature. Along with the annealing step, the extension step was at 72°C for the most effective time depending on the length of the target region. Also, all the Touchdown PCR amplifications ended with an extension at 72°C for 5 minutes and the PCR products were stored at 4°C before running on gels.

A positive control (containing ADP1 genomic DNA as the template) and a negative control (nuclease-free water as template) were included in each PCR amplification experiment.

Checking PCR products on gels

Two different percentages of agarose gels were used in the project. One was an 0.8% gel (30 ml volume) which was used to run the larger DNA product pieces (size range from 500 bp to 4000 bp) with its optimal DNA marker; the other one was a 1.3% gel (30 ml volume) which was used to run the smaller pieces of DNA products (size range from 100 bp to 1000 bp) with its optimal DNA marker. For each gel, 10 µl of the PCR product was run on the gel.

PCR products clean-up system

The Wizard SV gel and PCR clean-up system (Promega) was applied to clean up PCR products before cloning in order to remove the remainder of PCR reagents and to increase cloning efficiency.

Cloning and transformation

Cloning was performed using the pGEM-T Easy DNA vector (Promega). For each clone, 3µl of a PCR product was added to T4 ligase, 2x T4 ligase buffer, pGEMT-Easy vector, and nuclease-free water up to a 10 µl ligation volume. All the ligations were incubated at 4°C overnight.

On the next day, 2.5-5 µl of the ligation mixture was transformed into 50 µl DH5α competent *E. coli* (Invitrogen). The transformant was incubated on ice for 30 minutes. Then, the transformant was heat shocked for 20 seconds in a 37°C water bath without shaking and immediately placed on ice for 2 minutes. Subsequently, 950 µl of pre-warmed NB was added to each transformant. Then, the transformant was incubated at 37°C for 60 minutes at 225 rpm. After transformation was done, 1 ml liquid culture was pelleted, re-suspended into 200µl, and spread onto 2 ampicillin and X-Gal-containing Luria-Bertani (LB) plates, respectively. The culture plates were inverted at 35°C for 16 to 20 hours.

On the next day, the culture plates were moved from incubator to refrigerator for 2-3 hours. Then, several white colonies (usually, 2-5 white colonies) were selected

to re-inoculate to 3 ml ampicillin-containing nutrient broth, shaking culture at 180 rpm for 16 to 20 hours.

Plasmid purification

The QIAGEN Plasmid Purification kit (QIAGEN) was used to purify the plasmids from transformants. After purifications were done, the plasmid DNA were digested with restriction enzyme *Eco*RI (the insertion sites flanking the T-A cloning site of pGEMT-Easy DNA vector) in a 20 µl reaction and then incubated in 37°C water bath overnight. All the digestions were run on 0.8% gels to confirm the qualities and contents of the plasmids. Both the original purified plasmid and positive control from the original PCR products were also applied as size controls. The plasmid DNA was stored in the -20°C freezer.

Sequencing plasmid DNA

The sequencing of plasmid DNA was performed with PCR cycle sequencing, using fluorescently-labeled M13 Forward and Reverse primers. Sequencing via PCR amplification is slightly different than regular PCR amplification.

For each run of sequencing via PCR amplification, 1 µl of the plasmid template was mixed with 1.5 µl M13F and 1.5 µl M13R fluorescently-labeled universal primers, 7.2 µl of 3.5X SequiTherm EXCEL II Sequencing Buffer, and 1 µl of SequiTherm EXCEL II DNA Polymerase. Then, 17 µl of a total reaction volume

was distributed into 4 different PCR tubes, each of which contained 2 μl of dATP, dTTP, dCTP, dGTP, respectively.

The amplification was held at 92°C for 2 minutes. The examined plasmid template was denatured at 92°C for 30 seconds, annealed at 54°C for 30 seconds and extended at 72°C for 1 minute. The total cycles for each run were 30. The sequencing PCR products were then either stored at 4°C or added the stop solution directly. 3 μl of Stop/Loading Buffer was added to each tube containing sequencing PCR products. The mixture was then heated to 95°C for 3 minutes and all of tubes were moved onto ice immediately.

The DNA sequencing was done via applying 0.6 μl of each examined plasmid to an SDS-polyacrylamide gel on the LI-COR 4300S DNA Sequencer. After DNA sequencing was done, the sequenced results were interpreted and confirmed on both channels (channel 700 & 800), which correspond to both strands of DNA. Thus, both strands of DNA were sequenced for each template.

Analysis of sequence data

The analysis of sequence data was conducted via the following software: (1) NCBI – BLAST (Basic Local Alignment Search Tool; blastn, blastp, and tblastx); (2) the BCM search launcher – pairwise sequence alignment and multiple sequence alignment (http://searchlauncher.bcm.tmc.edu/); (3) Vector NTI Advance™ 10 (Invitrogen) – the AlignX program.

The analysis of sequence data focused on the comparison of nucleotides and amino acids in *Acinetobacter* strains to those in ADP1 or other microbes.

Results

The analyzed data of the *umuD* and *umuC* genes in other *Acinetobacter* strains were compared to those of ADP1 and eventually, *E. coli* and other bacteria. My objectives were to determine whether the *umuD* and *umuC* genes from *Acinetobacter* strains other than ADP1 (i) also have an elongated *umuD* and a nearly truncated *umuC* as those present in ADP1, or (ii) contain the same *umuDC* operon as does *E. coli*. I used Touchdown PCR amplification to amplify the *umuDC* homologs from *Acinetobacter* strains other than ADP1, cloned and sequenced these specific PCR products, and analyzed them for the presence of functional protein motifs conserved and required for *E. coli* UmuD function.

The results from regular PCR amplification

One set of primers (16S rRNA For and 16S rRNA Rev) was used to perform a quality control PCR experiment on DNA isolated from all of the examined strains, as the presence of 16S ribosomal RNA is conserved among *Acinetobacter* species. The predicted amplified size for this PCR product is 575 bp. DNA from all the strains used in this project successfully yielded a product of this size when amplified with this set of primers, demonstrating that the genomic DNA purified from *Acinetobacter* species strains were all suitable for PCR amplification (Table 3).

The examined strains were then run in regular PCR amplification experiments with numerous different sets of primers (Table 2 and 3). Unfortunately, except for the

Table 3. Primers used in either regular PCR amplification, Touchdown PCR amplification, or inverse PCR amplification.

Strain	Primer pair	16S rRNA	8 & 5	0 & 5	A & 9.2	0 & B	A & N
	Size	575 bp	292 bp	504 bp	634 bp	651 bp	1479 bp
	Target region	16S rRNA	*umuD*	*umuD*	*umuD*	*umuDC*	*umuDC*
Acinetobacter baylyi ADP1		p	p	p	p	p	p
Acinetobacter radioresistens		p	p	**p***	np	**p***	np
Acinetobacter genomospecies 3		p	p	**p***	np	p	np
Acinetobacter junii		p	p	**p***	np	np	np
Acinetobacter genomospecies 9		p	p	**p***	np	np	**p***
Acinetobacter anitratus		p	p	**p***	np	np	np
Acinetobacter johnsonii		p	**p***	**p***	**p***	np	np
Acinetobacter conjunctivae TU14		p	-	2p	np	**p***	p
Acinetobacter genomospecies 16		p	-	np	np	np	np
Acinetobacter haemolyticus		p	-	np	-	np	-
Acinetobacter genomospecies 6		p	-	np	-	np	-
Acinetobacter genomospecies 10		p	-	-	-	np	-
Acinetobacter calcoaceticus		p	-	-	-	np	-
Acinetobacter genomospecies 14		p	p	**p***	-	np	-
Escherichia coli		-	-	-	-	-	-

p* = cloned and analyzed PCR product.
p = PCR product observed of approximate correct size.
2p = 2 PCR products observed
np = no product.
- = not done

Table 3. Primers used in either regular PCR amplification, Touchdown PCR amplification, or inverse PCR amplification. (continued)

Strain	Primer pair	A & K	A & B	L & B	10.2 & N	2 & B	P & O
	Size	808 bp	918 bp	111 bp	845 bp	68 bp	315 bp
	Target region	*umuD*	*umuDC*	*umuC*	*umuC*	*umuC*	*umuC*
Acinetobacter baylyi ADP1		p	p	p	p	p	p
Acinetobacter radioresistens		-	np	np	np	-	np
Acinetobacter genomospecies 3		-	np	np	np	np	np
Acinetobacter junii		-	np	np	np	np	np
Acinetobacter genomospecies 9		-	np	np	np	np	np
Acinetobacter anitratus		-	np	np	np	np	np
Acinetobacter johnsonii		np	np	np	np	np	np
Acinetobacter conjunctivae TU14		np	p	np	np	-	np
Acinetobacter genomospecies 16		np	np	np	np	np	np
Acinetobacter haemolyticus		np	-	-	-	-	-
Acinetobacter genomospecies 6		np	-	-	-	-	-
Acinetobacter genomospecies 10		np	-	-	-	-	-
Acinetobacter calcoaceticus		np	-	-	-	-	-
Acinetobacter genomospecies 14		np	-	-	-	-	-
Escherichia coli		-	-	-	-	-	-

p* = cloned and analyzed PCR product.
p = PCR product observed of approximate correct size.
np = no product.
- = not done

Table 3. Primers used in either regular PCR amplification, Touchdown PCR amplification, or inverse PCR amplification. (continued)

Strain	Primer pair	**0 & N**	**8 & N**	**2 & N**	**10.2 & B**	From umuD & To umuD	**2 & 4**
	Size	1212 bp	~1000 bp	629 bp	284 bp	770 bp	1706 bp
	Target region	*umuDC*	*umuDC*	*umuC*	*umuC*	*umuD*	*umuC*
Acinetobacter baylyi ADP1		np	p	p	p	p	p
Acinetobacter radioresistens		np	np	np	-	np	np
Acinetobacter genomospecies 3		-	-	-	np	np	np
Acinetobacter junii		-	-	-	np	np	np
Acinetobacter genomospecies 9		-	-	-	-	np	np
Acinetobacter anitratus		-	-	-	-	np	np
Acinetobacter johnsonii		np	np	np	np	np	np
Acinetobacter conjunctivae TU14		np	np	np	-	-	np
Acinetobacter genomospecies 16		-	-	-	-	np	np
Acinetobacter haemolyticus		-	-	-	-	np	-
Acinetobacter genomospecies 6		-	-	-	-	np	-
Acinetobacter genomospecies 10		-	-	-	-	np	-
Acinetobacter calcoaceticus		-	-	-	-	np	-
Acinetobacter genomospecies 14		-	-	-	-	np	-
Escherichia coli		-	-	-	-	-	-

p* = cloned and analyzed PCR product.
p = PCR product observed of approximate correct size.
np = no product.
- = not done

Table 3. Primers used in either regular PCR amplification, Touchdown PCR amplification, or inverse PCR amplification. (continued)

Strain	Primer pair	3 & 2	10 & 4	10 & B	10 & 3	6 & 3	9.2 & 10.2 (inverse PCR)
	Size	1211 bp	645 bp	225 bp	1394 bp	1211 bp	unknown
	Target region	*umuC*	*umuC*	*umuC*	*umuC*	*umuC*	unknown
Acinetobacter baylyi ADP1		np	p	p	np	-	p
Acinetobacter radioresistens		np	-	-	-	-	np
Acinetobacter genomospecies 3		np	np	np	-	-	np
Acinetobacter junii		np	np	np	-	-	np
Acinetobacter genomospecies 9		np	-	np	-	-	-
Acinetobacter anitratus		np	np	np	-	-	-
Acinetobacter johnsonii		np	-	-	-	-	np
Acinetobacter conjunctivae TU14		np	np	np	np	-	-
Acinetobacter genomospecies 16		-	np	np	-	-	np
Acinetobacter haemolyticus		-	-	-	-	-	-
Acinetobacter genomospecies 6		-	-	-	-	-	-
Acinetobacter genomospecies 10		-	-	-	-	-	-
Acinetobacter calcoaceticus		-	-	-	-	-	-
Acinetobacter genomospecies 14		-	-	-	-	-	-
Escherichia coli		-	-	-	-	p	-

p* = cloned and analyzed PCR product.
p = PCR product observed of approximate correct size.
np = no product.
- = not done

positive control (a PCR reaction containing ADP1 genomic DNA), there were no PCR products amplified from the other *Acinetobacter* strains with these various sets of primers, which were designed from either ADP1 or *E. coli*. Therefore, a Touchdown PCR method was tried with the same primers because of its suitability for amplifying product from a template with potential mismatches to the primers.

Touchdown PCR amplification

The annealing between a primer and template is an important factor that affects the result of PCR amplification. When the annealing temperature is above the primer's T_m, it will yield no product. On the other hand, when the temperature is too far below the T_m, promiscuous primer binding at ectopic sites will give unwanted products. When PCR is attempted across species boundaries, the optimal T_m may not be able to be precisely calculated due to the likely sequence differences among different strains. An alternative PCR method, Touchdown PCR amplification, was developed to overcome much of the uncertainty in creating an optimized PCR protocol (Don *et al.*, 1991). Touchdown PCR is a PCR amplification method that uses non-degenerate primer mismatches to the templates, which can yield homogeneous PCR products (Roux, 1994).

In the protocol, "touchdown" refers to the gradual lowering of annealing temperature to facilitate the eventual binding of primer to the imperfectly matching template. The annealing temperature was decreased by 1°C for every 3 cycles of the PCR amplification and followed by 10 additional cycles at the lowest annealing

temperature. The total cycle number for each Touchdown PCR in this project was 43 cycles. Numerous sets of ADP1-specific primers were used in Touchdown PCR (Table 3), whereupon certain target regions were amplified with some sets of these primers. Also, positive and negative controls were always done when Touchdown PCR was applied.

Analyses of DNA sequence data

The method of Touchdown PCR amplified the regions of interest from some but not all *Acinetobacter* strains tested. After cloning PCR products into the pGEMT-T Easy Vector, plasmid DNA was transformed into and harvested from *E.coli* DH5α cells. Several plasmids were sent to companies and sequenced, but most were sequenced onsite with the LI-COR DNA Sequencer. All the DNA sequence data were primarily analyzed via NCBI Nucleotide-nucleotide BLAST (blastn) and Translated query vs. protein database (blastx). The nucleotide-nucleotide BLAST (blastn) searches if there is any nucleotide homolog from the NCBI database matching to the query nucleotide sequence. The translated query vs. protein database (blastx) can translate the input nucleotide sequence to an amino acid sequence and search if there is any known, functional protein, or putative protein, which matches to the encoded query sequences from NCBI database. After obtaining the results from the NCBI BLAST programs, the gene homologs and gene encoded homologs of the test strains were aligned with the nucleotide sequences or protein sequences from other bacterial model systems via using Vector NTI program.

Acinetobacter sp. genomospecies 9

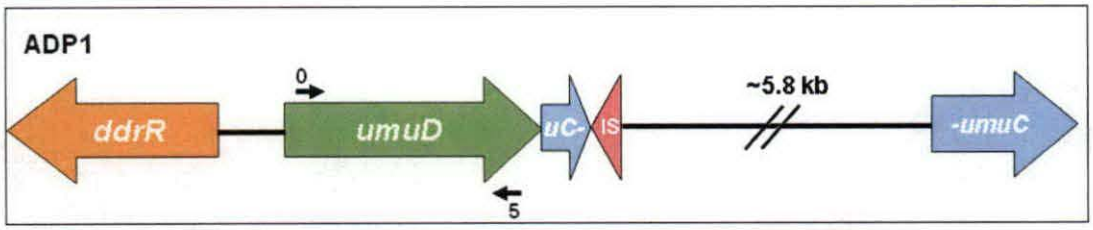
The sequencing result from the region (~800 bp) which was amplified with primers A and N in *Acinetobacter* genomospecies 9 encoded a tranposase-like protein that is present in various *Acinetobacter* species. The best match protein was *Acinetobacter haemolyticus* sequence IS17 putative transposase. However, according to the sequencing result, this amplified region was somewhere on chromosome rather than the predicted target region which contains the *umuDC* homolog.

umuD homologs are relatively conserved in examined Acinetobacter strains

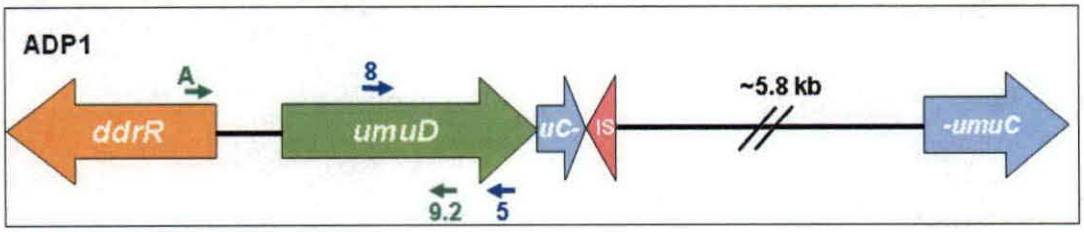
The primers 0 and 5 amplify nearly the whole region of the *umuD* gene (504 bp out of 612 bp) in ADP1 (Table 3 and Figure 4a). Except for the positive control ADP1, no other examined strains were successfully amplified by the primers 0 and 5 via the regular PCR method. Five examined strains were amplified by this pair of primers via the Touchdown PCR method. Although there were eight strains found containing the *umuD* homolog, three out of eight strains were amplified by other pairs of primers. The positive (ADP1) and negative (no DNA) controls were always performed along with the examined strains when Touchdown PCR methods were executed.

At the nucleotide level, the *umuD* homolog in *A. junii*, *A.* genomospecies 9, and *A.* genomospecies 14 were 100% identical to that in ADP1 (Figure 5) and their encoded UmuD homologs also had 100% identity (167/167 amino acids) (Figure 6).

(a)



(b)



(c)

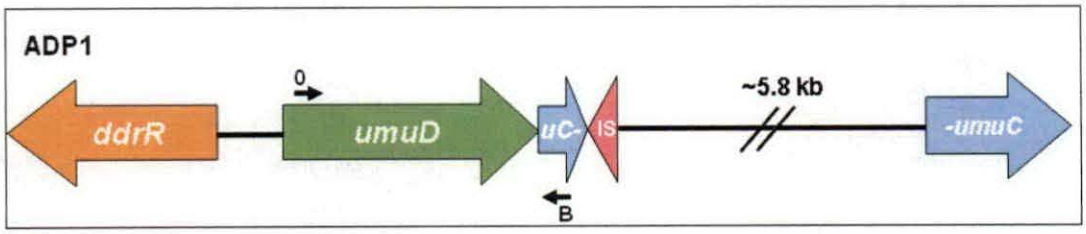
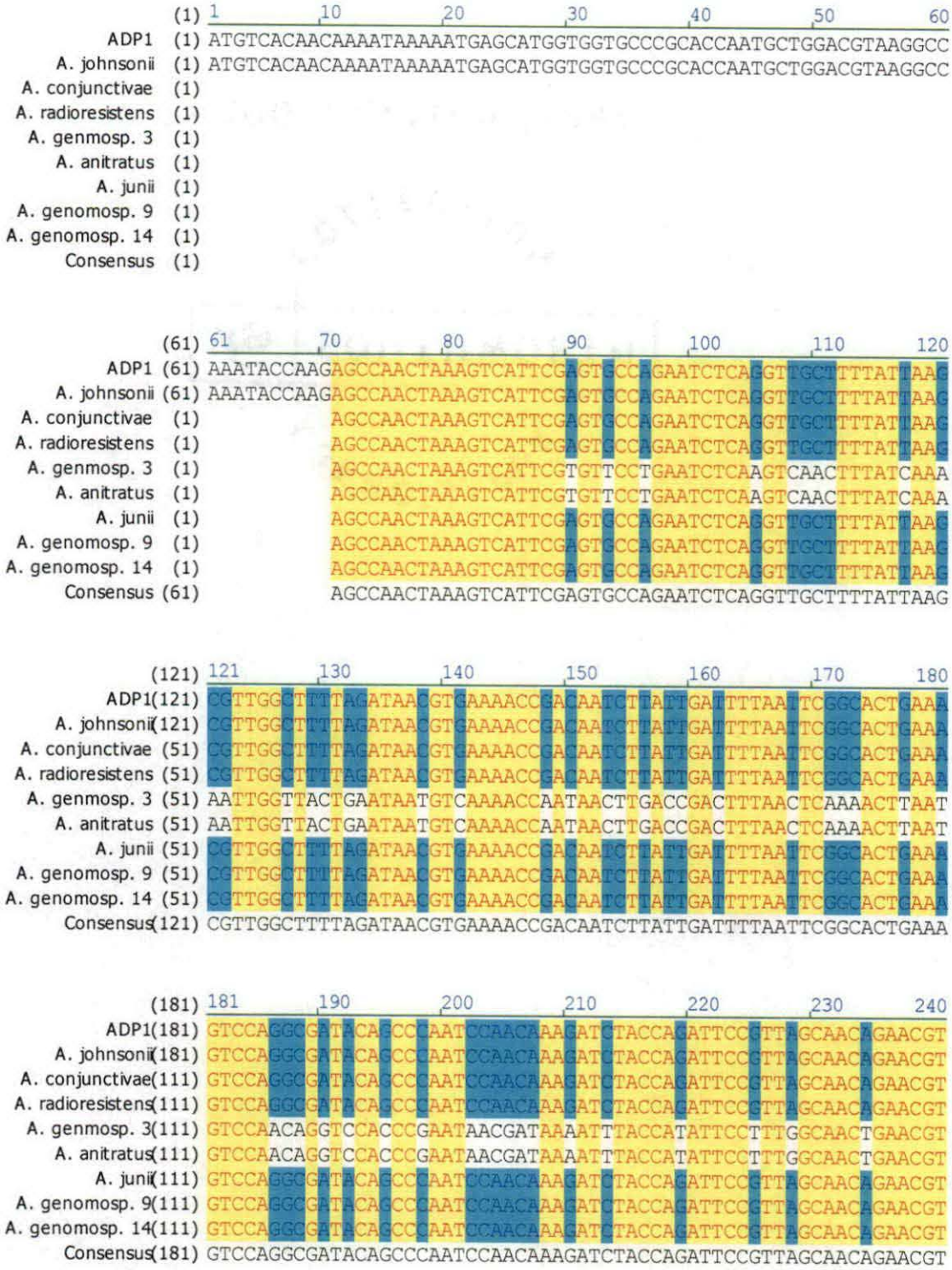
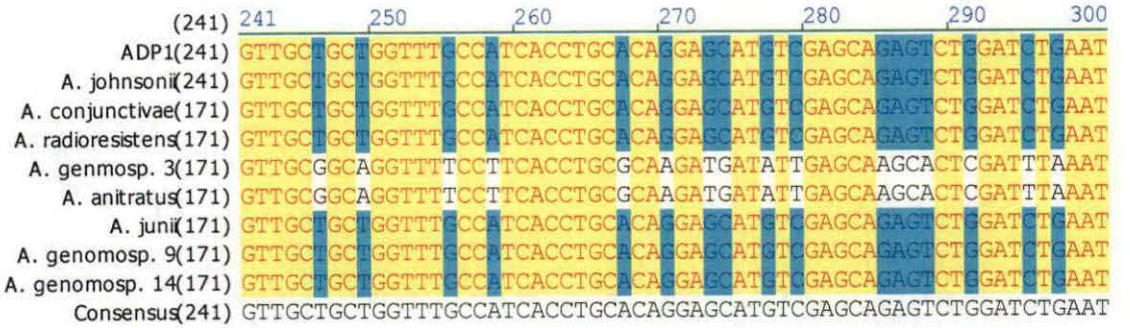
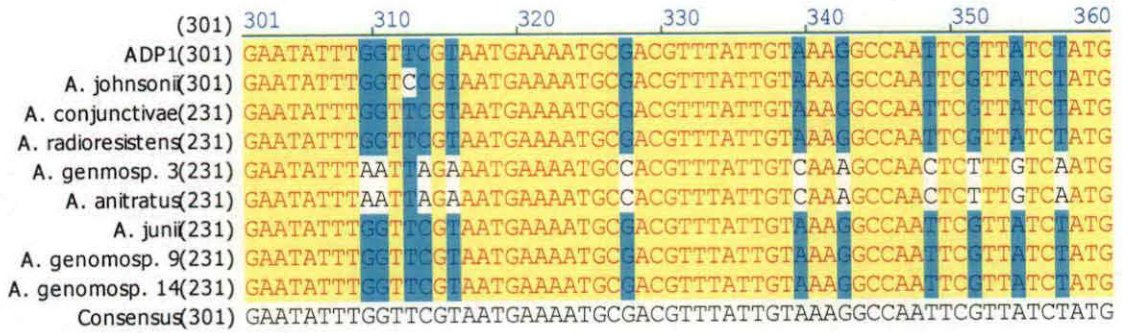
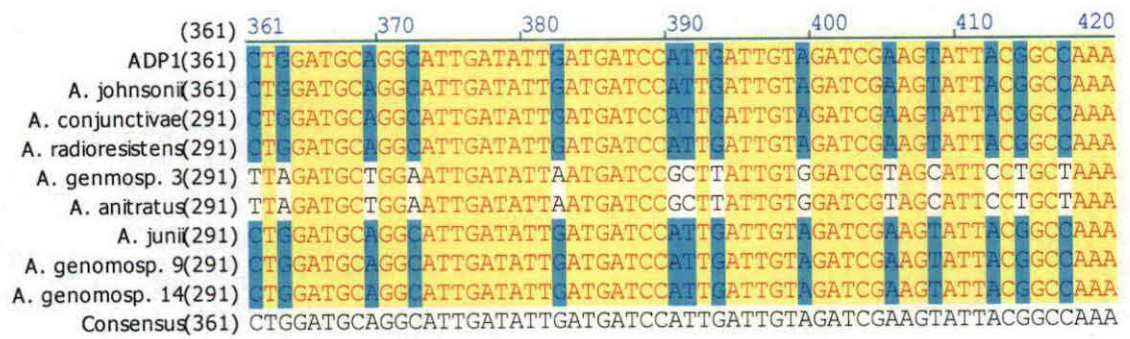


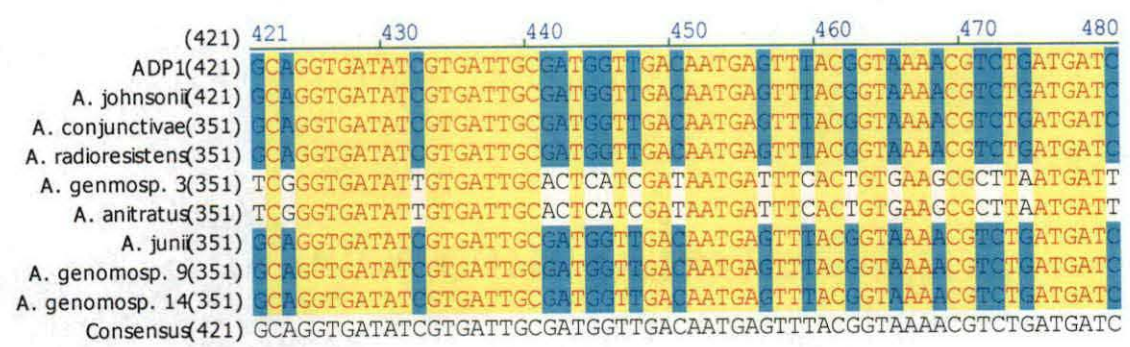
Figure 4. The binding sites for the pairs of primers on the target regions. (a) The pair of primers 0 and 5 amplified the target region (504 bp) of the *umuD* homolog. (b) Two pairs of primers 8-5 (292 bp) and A-9.2 (634bp) amplified the overlapping target region (770 bp) of the *ddrR* homolog, *ddrR-umuD* promoter region, and the *umuD* homolog. (c) The pair of primers 0 and B amplified the target region (651 bp) of the *umuD* homolog and partial small *umuC* fragment.


(1) 1 10 20 30 40 50 60
ADP1 (1) ATGTCACAACAAAATAAAAATGAGCATGGTGGTGCCCGCACCAATGCTGGACGTAAGGCC
A. johnsonii (1) ATGTCACAACAAAATAAAAATGAGCATGGTGGTGCCCGCACCAATGCTGGACGTAAGGCC
A. conjunctivae (1)
A. radioresistens (1)
A. genmosp. 3 (1)
A. anitratus (1)
A. junii (1)
A. genomosp. 9 (1)
A. genomosp. 14 (1)
Consensus (1)
(61) 61 70 80 90 100 110 120
ADP1 (61) AAATACCAAGAGCCAACTAAAGTCATTCGAGTGCCAGAATCTCAGGTTGCTTTTATTAAG
A. johnsonii (61) AAATACCAAGAGCCAACTAAAGTCATTCGAGTGCCAGAATCTCAGGTTGCTTTTATTAAG
A. conjunctivae (1) AGCCAACTAAAGTCATTCGAGTGCCAGAATCTCAGGTTGCTTTTATTAAG
A. radioresistens (1) AGCCAACTAAAGTCATTCGAGTGCCAGAATCTCAGGTTGCTTTTATTAAG
A. genmosp. 3 (1) AGCCAACTAAAGTCATTCGTGTTCCTGAATCTCAAGTCAACTTTATCAAA
A. anitratus (1) AGCCAACTAAAGTCATTCGTGTTCCTGAATCTCAAGTCAACTTTATCAAA
A. junii (1) AGCCAACTAAAGTCATTCGAGTGCCAGAATCTCAGGTTGCTTTTATTAAG
A. genomosp. 9 (1) AGCCAACTAAAGTCATTCGAGTGCCAGAATCTCAGGTTGCTTTTATTAAG
A. genomosp. 14 (1) AGCCAACTAAAGTCATTCGAGTGCCAGAATCTCAGGTTGCTTTTATTAAG
Consensus (61) AGCCAACTAAAGTCATTCGAGTGCCAGAATCTCAGGTTGCTTTTATTAAG
(121) 121 130 140 150 160 170 180
ADP1(121) CGTTGGCTTTTAGATAACGTGAAAACCGACAATCTTATTGATTTTAATTCGGCACTGAAA
A. johnsonii(121) CGTTGGCTTTTAGATAACGTGAAAACCGACAATCTTATTGATTTTAATTCGGCACTGAAA
A. conjunctivae (51) CGTTGGCTTTTAGATAACGTGAAAACCGACAATCTTATTGATTTTAATTCGGCACTGAAA
A. radioresistens (51) CGTTGGCTTTTAGATAACGTGAAAACCGACAATCTTATTGATTTTAATTCGGCACTGAAA
A. genmosp. 3 (51) AATTGGTTACTGAATAATGTCAAAACCAATAACTTGACCGACTTTAACTCAAAACTTAAT
A. anitratus (51) AATTGGTTACTGAATAATGTCAAAACCAATAACTTGACCGACTTTAACTCAAAACTTAAT
A. junii (51) CGTTGGCTTTTAGATAACGTGAAAACCGACAATCTTATTGATTTTAATTCGGCACTGAAA
A. genomosp. 9 (51) CGTTGGCTTTTAGATAACGTGAAAACCGACAATCTTATTGATTTTAATTCGGCACTGAAA
A. genomosp. 14 (51) CGTTGGCTTTTAGATAACGTGAAAACCGACAATCTTATTGATTTTAATTCGGCACTGAAA
Consensus(121) CGTTGGCTTTTAGATAACGTGAAAACCGACAATCTTATTGATTTTAATTCGGCACTGAAA
(181) 181 190 200 210 220 230 240
ADP1(181) GTCCAGGCGATACAGCCCAATCCAACAAAGATCTACCAGATTCCGTTAGCAACAGAACGT
A. johnsonii(181) GTCCAGGCGATACAGCCCAATCCAACAAAGATCTACCAGATTCCGTTAGCAACAGAACGT
A. conjunctivae(111) GTCCAGGCGATACAGCCCAATCCAACAAAGATCTACCAGATTCCGTTAGCAACAGAACGT
A. radioresistens(111) GTCCAGGCGATACAGCCCAATCCAACAAAGATCTACCAGATTCCGTTAGCAACAGAACGT
A. genmosp. 3(111) GTCCAACAGGTCCACCCGAATAACGATAAAATTTACCATATTCCTTTGGCAACTGAACGT
A. anitratus(111) GTCCAACAGGTCCACCCGAATAACGATAAAATTTACCATATTCCTTTGGCAACTGAACGT
A. junii(111) GTCCAGGCGATACAGCCCAATCCAACAAAGATCTACCAGATTCCGTTAGCAACAGAACGT
A. genomosp. 9(111) GTCCAGGCGATACAGCCCAATCCAACAAAGATCTACCAGATTCCGTTAGCAACAGAACGT
A. genomosp. 14(111) GTCCAGGCGATACAGCCCAATCCAACAAAGATCTACCAGATTCCGTTAGCAACAGAACGT
Consensus(181) GTCCAGGCGATACAGCCCAATCCAACAAAGATCTACCAGATTCCGTTAGCAACAGAACGT


(241) 241 250 260 270 280 290 300
ADP1(241) GTTGCTGCTGGTTTGCCATCACCTGCACAGGAGCATGTCGAGCAGAGTCTGGATCTGAAT
A. johnsonii(241) GTTGCTGCTGGTTTGCCATCACCTGCACAGGAGCATGTCGAGCAGAGTCTGGATCTGAAT
A. conjunctivae(171) GTTGCTGCTGGTTTGCCATCACCTGCACAGGAGCATGTCGAGCAGAGTCTGGATCTGAAT
A. radioresistens(171) GTTGCTGCTGGTTTGCCATCACCTGCACAGGAGCATGTCGAGCAGAGTCTGGATCTGAAT
A. genmosp. 3(171) GTTGCGGCAGGTTTTCCTTCACCTGCGCAAGATGATATTGAGCAAGCACTCGATTTAAAT
A. anitratus(171) GTTGCGGCAGGTTTTCCTTCACCTGCGCAAGATGATATTGAGCAAGCACTCGATTTAAAT
A. junii(171) GTTGCTGCTGGTTTGCCATCACCTGCACAGGAGCATGTCGAGCAGAGTCTGGATCTGAAT
A. genomosp. 9(171) GTTGCTGCTGGTTTGCCATCACCTGCACAGGAGCATGTCGAGCAGAGTCTGGATCTGAAT
A. genomosp. 14(171) GTTGCTGCTGGTTTGCCATCACCTGCACAGGAGCATGTCGAGCAGAGTCTGGATCTGAAT
Consensus(241) GTTGCTGCTGGTTTGCCATCACCTGCACAGGAGCATGTCGAGCAGAGTCTGGATCTGAAT


(301) 301 310 320 330 340 350 360
ADP1(301) GAATATTTGGTTCGTAATGAAAATGCGACGTTTATTGTAAAGGCCAATTCGTTATCTATG
A. johnsonii(301) GAATATTTGGTCCGTAATGAAAATGCGACGTTTATTGTAAAGGCCAATTCGTTATCTATG
A. conjunctivae(231) GAATATTTGGTTCGTAATGAAAATGCGACGTTTATTGTAAAGGCCAATTCGTTATCTATG
A. radioresistens(231) GAATATTTGGTTCGTAATGAAAATGCGACGTTTATTGTAAAGGCCAATTCGTTATCTATG
A. genmosp. 3(231) GAATATTTAATTAGAAATGAAAATGCCACGTTTATTGTCAAAGCCAACTCTTTGTCAATG
A. anitratus(231) GAATATTTAATTAGAAATGAAAATGCCACGTTTATTGTCAAAGCCAACTCTTTGTCAATG
A. junii(231) GAATATTTGGTTCGTAATGAAAATGCGACGTTTATTGTAAAGGCCAATTCGTTATCTATG
A. genomosp. 9(231) GAATATTTGGTTCGTAATGAAAATGCGACGTTTATTGTAAAGGCCAATTCGTTATCTATG
A. genomosp. 14(231) GAATATTTGGTTCGTAATGAAAATGCGACGTTTATTGTAAAGGCCAATTCGTTATCTATG
Consensus(301) GAATATTTGGTTCGTAATGAAAATGCGACGTTTATTGTAAAGGCCAATTCGTTATCTATG


(361) 361 370 380 390 400 410 420
ADP1(361) CTGGATGCAGGCATTGATATTGATGATCCATTGATTGTAGATCGAAGTATTACGGCCAAA
A. johnsonii(361) CTGGATGCAGGCATTGATATTGATGATCCATTGATTGTAGATCGAAGTATTACGGCCAAA
A. conjunctivae(291) CTGGATGCAGGCATTGATATTGATGATCCATTGATTGTAGATCGAAGTATTACGGCCAAA
A. radioresistens(291) CTGGATGCAGGCATTGATATTGATGATCCATTGATTGTAGATCGAAGTATTACGGCCAAA
A. genmosp. 3(291) TTAGATGCTGGAATTGATATTAATGATCCGCTTATTGTGGATCGTAGCATTCCTGCTAAA
A. anitratus(291) TTAGATGCTGGAATTGATATTAATGATCCGCTTATTGTGGATCGTAGCATTCCTGCTAAA
A. junii(291) CTGGATGCAGGCATTGATATTGATGATCCATTGATTGTAGATCGAAGTATTACGGCCAAA
A. genomosp. 9(291) CTGGATGCAGGCATTGATATTGATGATCCATTGATTGTAGATCGAAGTATTACGGCCAAA
A. genomosp. 14(291) CTGGATGCAGGCATTGATATTGATGATCCATTGATTGTAGATCGAAGTATTACGGCCAAA
Consensus(361) CTGGATGCAGGCATTGATATTGATGATCCATTGATTGTAGATCGAAGTATTACGGCCAAA


(421) 421 430 440 450 460 470 480
ADP1(421) GCAGGTGATATCGTGATTGCGATGGTTGACAATGAGTTTACGGTAAAACGTCTGATGATC
A. johnsonii(421) GCAGGTGATATCGTGATTGCGATGGTTGACAATGAGTTTACGGTAAAACGTCTGATGATC
A. conjunctivae(351) GCAGGTGATATCGTGATTGCGATGGTTGACAATGAGTTTACGGTAAAACGTCTGATGATC
A. radioresistens(351) GCAGGTGATATCGTGATTGCGATGGTTGACAATGAGTTTACGGTAAAACGTCTGATGATC
A. genmosp. 3(351) TCGGGTGATATTGTGATTGCACTCATCGATAATGATTTCACTGTGAAGCGCTTAATGATT
A. anitratus(351) TCGGGTGATATTGTGATTGCACTCATCGATAATGATTTCACTGTGAAGCGCTTAATGATT
A. junii(351) GCAGGTGATATCGTGATTGCGATGGTTGACAATGAGTTTACGGTAAAACGTCTGATGATC
A. genomosp. 9(351) GCAGGTGATATCGTGATTGCGATGGTTGACAATGAGTTTACGGTAAAACGTCTGATGATC
A. genomosp. 14(351) GCAGGTGATATCGTGATTGCGATGGTTGACAATGAGTTTACGGTAAAACGTCTGATGATC
Consensus(421) GCAGGTGATATCGTGATTGCGATGGTTGACAATGAGTTTACGGTAAAACGTCTGATGATC

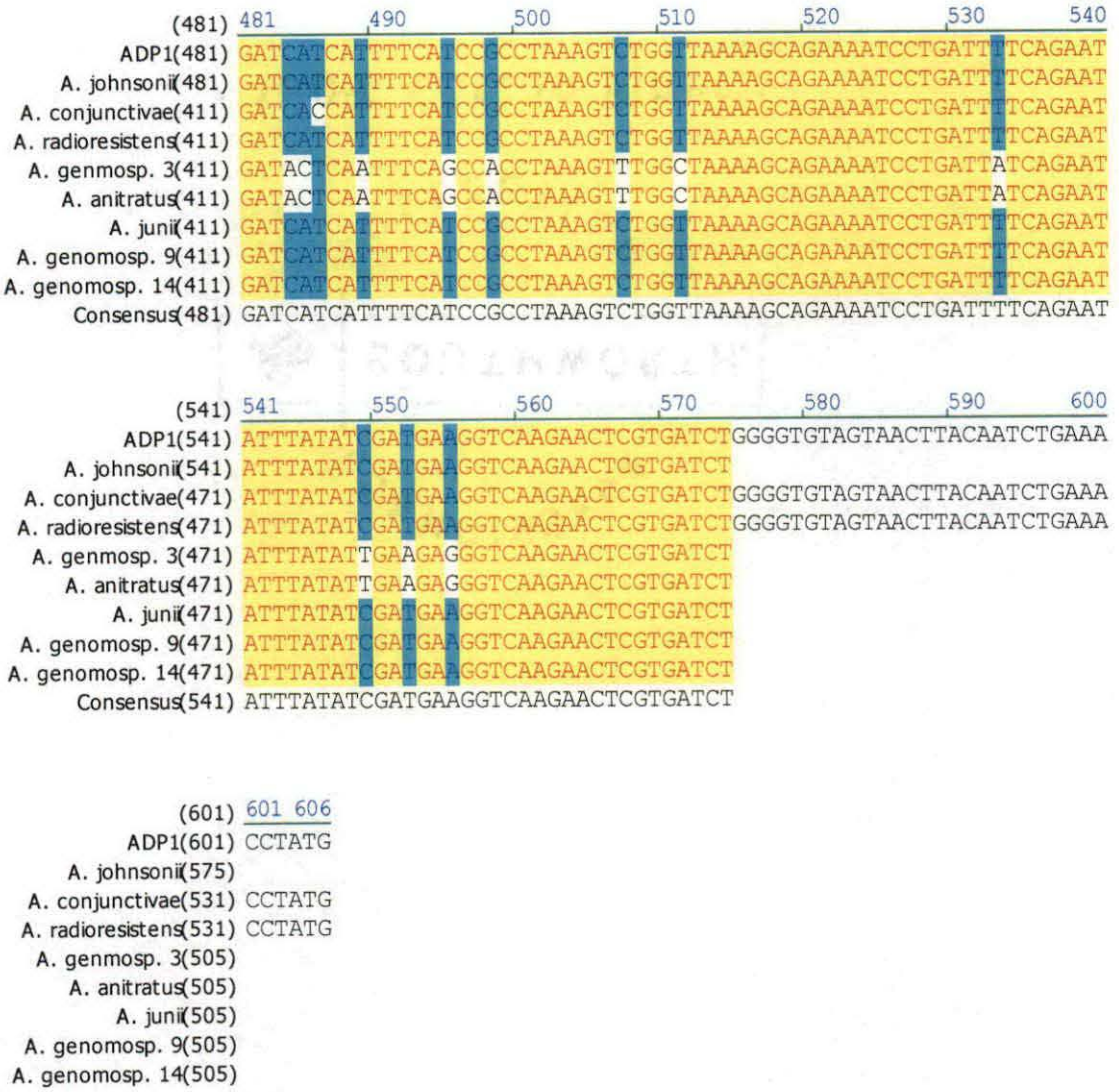


Figure 5: Nucleotide alignment of ADP1 and examined *Acinetobacter* strains. *umuD* homologs from eight sequenced strains were aligned with the *umuD* homolog in ADP1. For the discussion of *A. johnsonii*, *A. radioresistens*, and *A. conjunctivae* TU14, see next section.

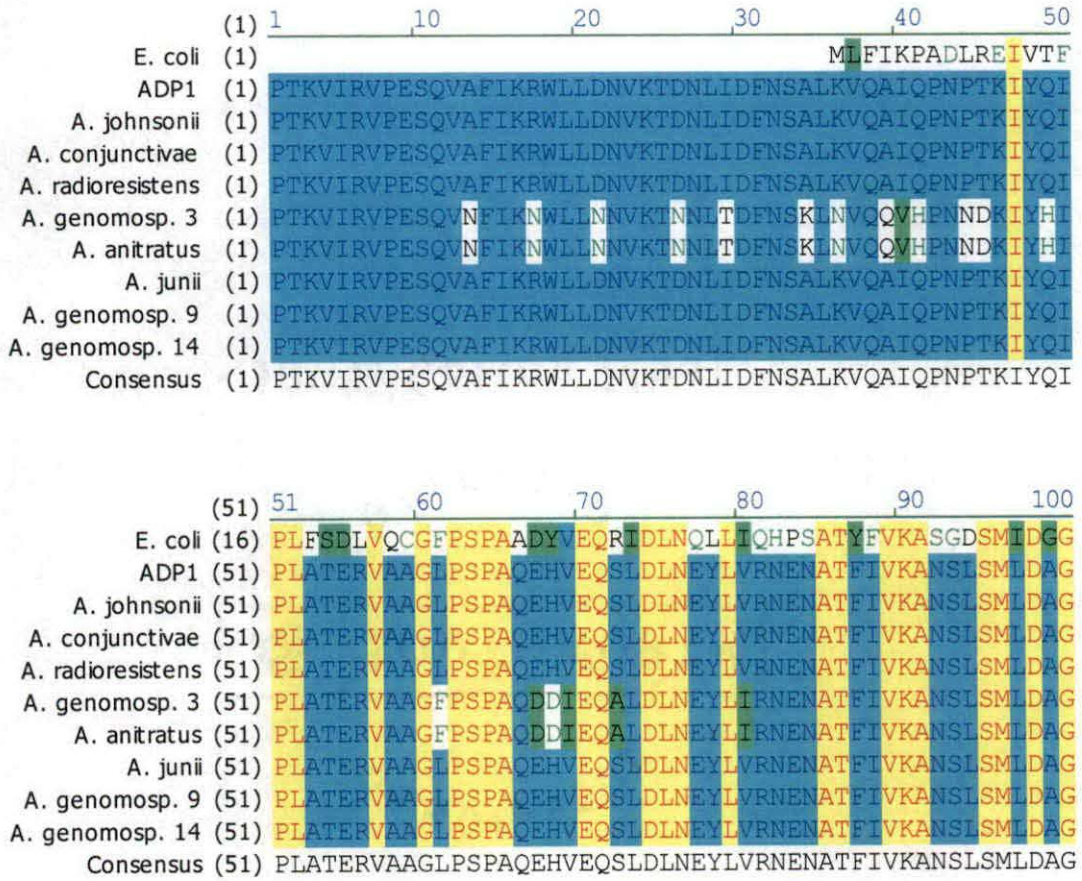


Figure 6. Amino acid alignment of the encoded UmuD homolog present in ADP1 and *Acinetobacter* strains. The yellow highlight corresponds to identical residues in all strains, and the blue high-light indicates a consensus residue not conserved in all strains.

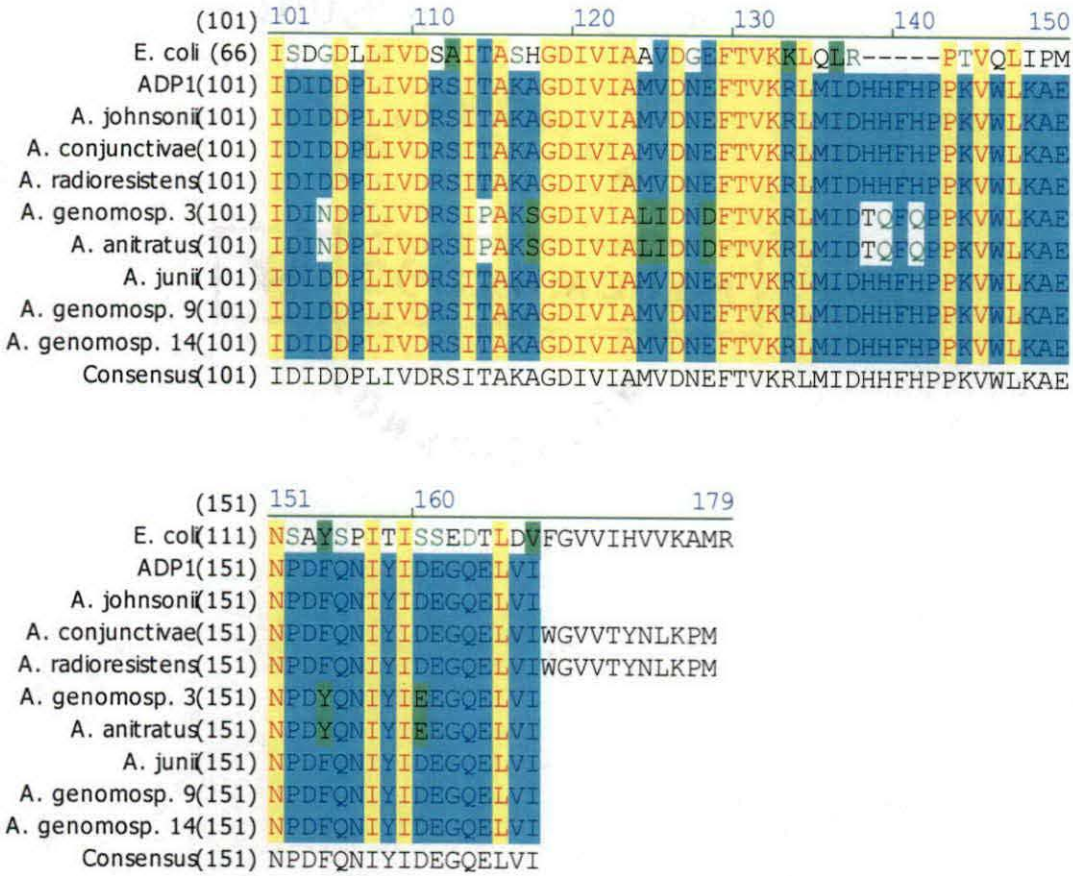


Figure 6. Amino acid alignment of the encoded UmuD homolog present in ADP1 and *Acinetobacter* strains. (continued) The yellow highlight corresponds to identical residues in all strains, and the blue highlight indicates a consensus residue not conserved in all strains.

On the other hand, the DNA sequences of the *umuD* homolog in *A.* genomospecies 3 and *A. anitratus* had a lower identity; only 76% nucleotide identity was shown in 504 residues (Figure 5). Therefore, these two strains had a slightly lower amino acid identity (137/168 amino acids; 81%), and similarity (152/168 amino acids; 90%) to the UmuD protein of ADP1 (Figure 6).

All of these five examined strains, *A. junii*, *A.* genomospecies 9, *A.* genomospecies 14, *A.* genomospecies 3 and *A. anitratus*, contain an alanine-24/glycine-25 putative cleavage site where E. coli self-cleaves. The catalytically active residues, serine-60 and lysine-97, required for self-cleavage in *E. coli* were also encoded by their *umuD* homologs.

The amino acids, leucine-101 and arginine-102, which are required for efficient UmuD self-cleavage in *E. coli,* are replaced by isoleucine-163 and aspartate-164 in ADP1. This change is also observed in the 5 examined strains. Moreover, the additional five amino acids (HHFHP) uniquely found in ADP1 which are adjacent to aspartate-164 were also found in *A. junii*, *A.* genomospecies 9 and *A.* genomospecies 14. Even though *A.* genomospecies 3 and *A. anitratus* also had these additional five amino acids, three out of five amino acids were changed; they were TQFQP rather than HHFHP.

Both a *ddrR* homolog and *umuD* homolog were present in *A. johnsonii*

The set of primers A and 9.2 (634 bp) and the set of primers 5 and 8 (292 bp) were used in an experiment designed to amplify the region from the *ddrR* gene

through the middle of the *umuD* gene in ADP1 (Figure 4b). The primers 5 and 8 (292 bp) amplify from the middle region of *umuD* to nearly the 3' end of *umuD* gene in ADP1. Taking the two sequences together, the sequenced region from primer A to primer 5 was 770 bp, resulting a short overlapping sequence between these two sets of primers. The examined strain, *A. johnsonii*, yielded appropriate sized products when using these two sets of primers by the Touchdown PCR method. The sequencing of these products showed that the *A. johnsonii* sequence had a very high identity (identity = 768/770 (99%), without any gap) to the region in ADP1. The sequenced region included the *ddrR* homolog, the *umuDC/ddrR* promoter region, and the *umuD* homolog.

Even though the actual size of *ddrR* in ADP1 is ~500 bp, the fragment of the *ddrR* homolog sequenced from *A. johnsonii* only contained 52 bp due to the location of the primer binding site. These 52 bp were 100% identical to that in ADP1.

144 bp were found in the *ddrR* and *umuD* promoter region. This region contains the putative -35 and -10 promoters for both *ddrR* and *umuD*, and their ribosome binding sites (Hare *et al.*, 2006). In the promoter region of *A. johnsonii*, two nucleotides were changed: the nucleotide A at 2674645 changed to G and the nucleotide at T 2674577 changed to A (Figure 7). However, the changes did not interrupt the putative promoters.

Even though the *umuD* homolog in *A. johnsonii* was 99.8% identical (573/574 nucleotides) to that region in ADP1 (Figure 5), the encoded amino acids were 100% identical (Figure 6), due to 1 silent mutation.

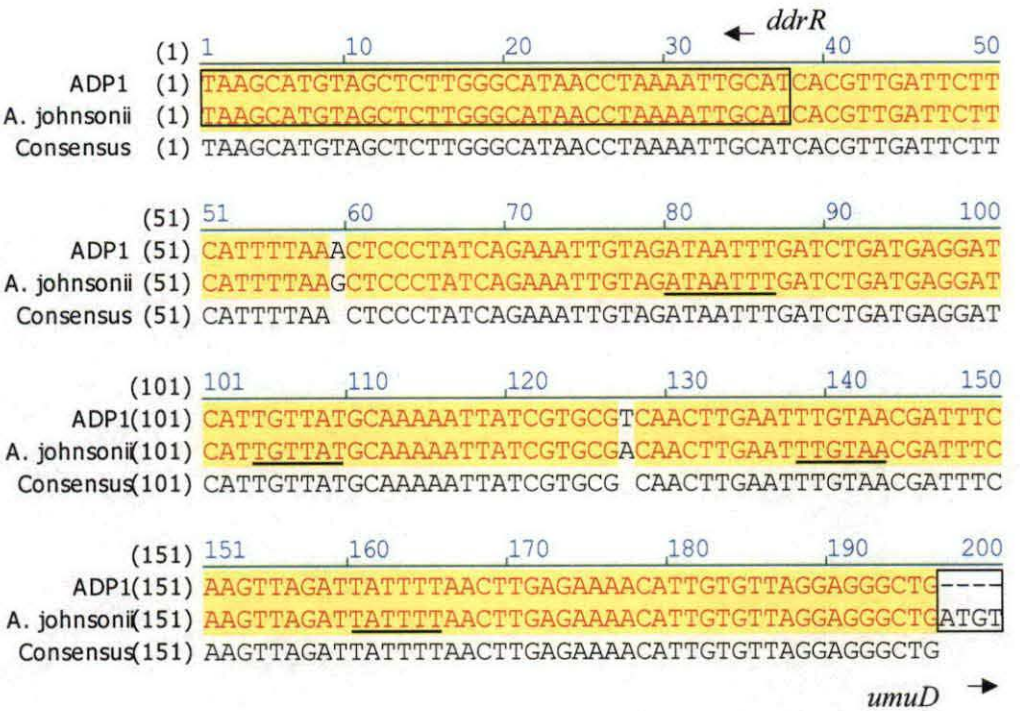


Figure 7. Nucleotide alignment of the *ddrR* homolog and the *ddrR–umuD* promoter region in ADP1 and *A. johnsonii*. Two gene homologs, *ddrR* and *umuD*, which orient oppositely are enclosed in boxes. The underlined nucleotides are the predicted promoter regions, -10 and -35. Moreover, the nucleotide change in *A. johnsonii* does not seem to disrupt the potential promoters.

A. johnsonii thus also contained a alanine-24/glycine-25 putative cleavage site corresponding to the site that is required for self-cleavage in *E. coli*. The putative catalytic residues, serine-60 and lysine-97, were also contained in its *umuD* homolog.

The amino acids, leucine-101 and arginine-102, which are required for efficient UmuD self-cleavage in *E. coli,* are replaced by isoleucine-163 and aspartate-164 in ADP1. This change was also observed in *A. johnsonii*. Moreover, the additional five amino acids (HHFHP) which are adjacent to aspartate-164 were also found in *A. johnsonii*.

Partial umuC homologs were present in two Acinetobacter strains

The set of primers 0 and B (651 bp) was used in another experiment (Figure 4c). The primers 0 and B are able to amplify both *umuD* and the first 39-a.a.-encoding-*umuC* fragment in ADP1. Except for the positive control ADP1, no other examined strains yielded PCR products with the primers 0 and B via regular PCR. The *umuDC* homolog region in *A. radioresistens* and *A. conjunctivae* strain TU14 were, however, amplified by this pair of primers via the Touchdown PCR method.

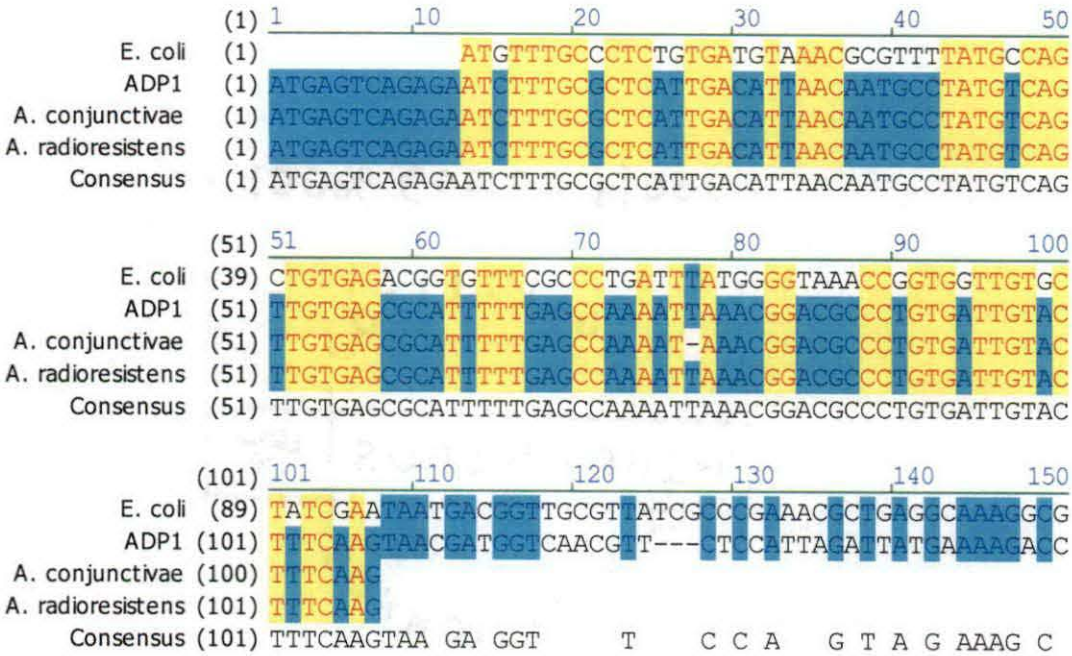
At the nucleotide level, the *umuD* and *umuC* homolog regions in *A. radioresistens* were 100% identical (651/651 nt.) to these two regions in ADP1 (Figures 5 and 8), thus it also had 100% amino acid identity (178/178 a.a) to the UmuD of ADP1 (Figure 6). Moreover, it also had 100% identity (35/35) to the UmuC of ADP1.

On the other hand, the DNA sequences of *umuDC* homologs in *A. conjunctivae* TU14 had a slightly lower identity to those in ADP1. The identity was 649/651 (99%), with one gap (Figures 5 and 8). The change of nucleotide was the nucleotide T at 2674022 changed to C. Even though this nucleotide was changed, there was no change at the amino acid level relative to ADP1. This kind of silent mutation is due to the degeneration of the DNA code (Figure 6).

One gap was found in the middle region of *umuC* homolog at which one T at 2673818 was omitted. A nonsense mutation was formed and as a consequence (Figure 8), the amino acid was changed from Leu (UUA) to a stop codon (UAA).

Both *A. radioresistens* and *A. conjunctivae* TU14 contain a alanine-24/glycine-25 putative cleavage site where is the site that is required for self-cleaved in *E. coli*. The putative catalytic residues, serine-60 and lysine-97, were also contained in their *umuD* homolog regions. These strains also had isoleucine-163 and aspartate-164 as those found in ADP1. Furthermore, the additional five amino acids (HHFHP) which are adjacent to aspartate-164 were also found in these two strains.

(A)



(B)

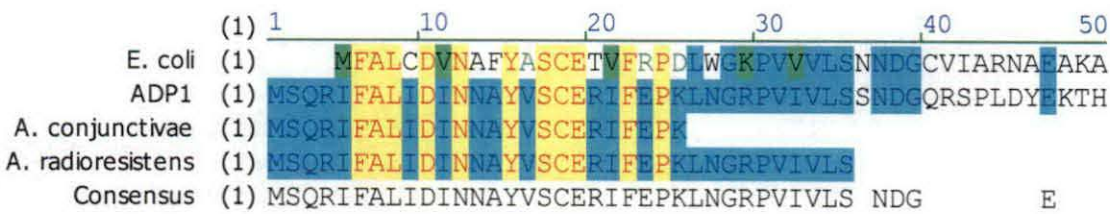


Figure 8. Alignment of the *umuC* homologs in *E. coli*, ADP1, *A. conjunctivae* strain TU14 and *A. radioresistens*. (A) Alignment of the *umuC* homolog. A gap is found in the nucleotide sequence in *A. conjunctivae*. (B) Alignment of the UmuC homolog. Only the first 39 amino acids in ADP1 are homologous to those in *E. coli*. The primers used in this experiment could only amplify the first 35 amino acids in both *A. radioresistens* and *A. conjunctivae*. However, since the gap is found in the two-thirds of the way through the encoding region, a leucine is changed to a stop codon in *A. conjuctivae*.

Putative functional protein motifs in UmuD

In order to compare the functional motifs from the extra-long UmuD protein found in ADP1 and other *Acinetobacter* strains to both normal-sized UmuD and other extra-long UmuD homologs, a multiple alignment of UmuD amino acid sequences was created using the Vector NTI program (Figure 9). The aligned UmuD were from bacteria, including the UmuD of *E. coli*, other normal-sized UmuDs in *Prochlorococcus, Vibrio, Citrobacter, Providencia,* and the extra-long UmuDs in ADP1, *Acinetobacter* species, *Thiomicrospira, Legionella, Synechococcus,* and *Chromobacterium.*

On average, the extra-long UmuD homologs possess ~60 more amino acids than those of normal-sized UmuD (Figure 9). A complete amino-terminus of UmuD starting at methionine was only confirmed in one strain, *A. johnsonii,* because its amplified target region contained the *ddrR* homolog, and *ddrR-umuD* promoter region upstream.

The recognition site for proteases ClpXP were not similar or conserved (relative to *E. coli*) in the all the aligned strains. In contrast to the ClpXP recognition site, the primary and secondary Lon recognition sites were highly conserved in *Acinetobacter* and all other bacteria, especially in the most important residue, proline, which is responsible for the structural folding of the UmuD homodimer amino-terminus (Gonzalez *et al.*, 2002). Even though the residues in both Lon sites were not 100% identical, the non-polar nature of these amino acids was not changed.

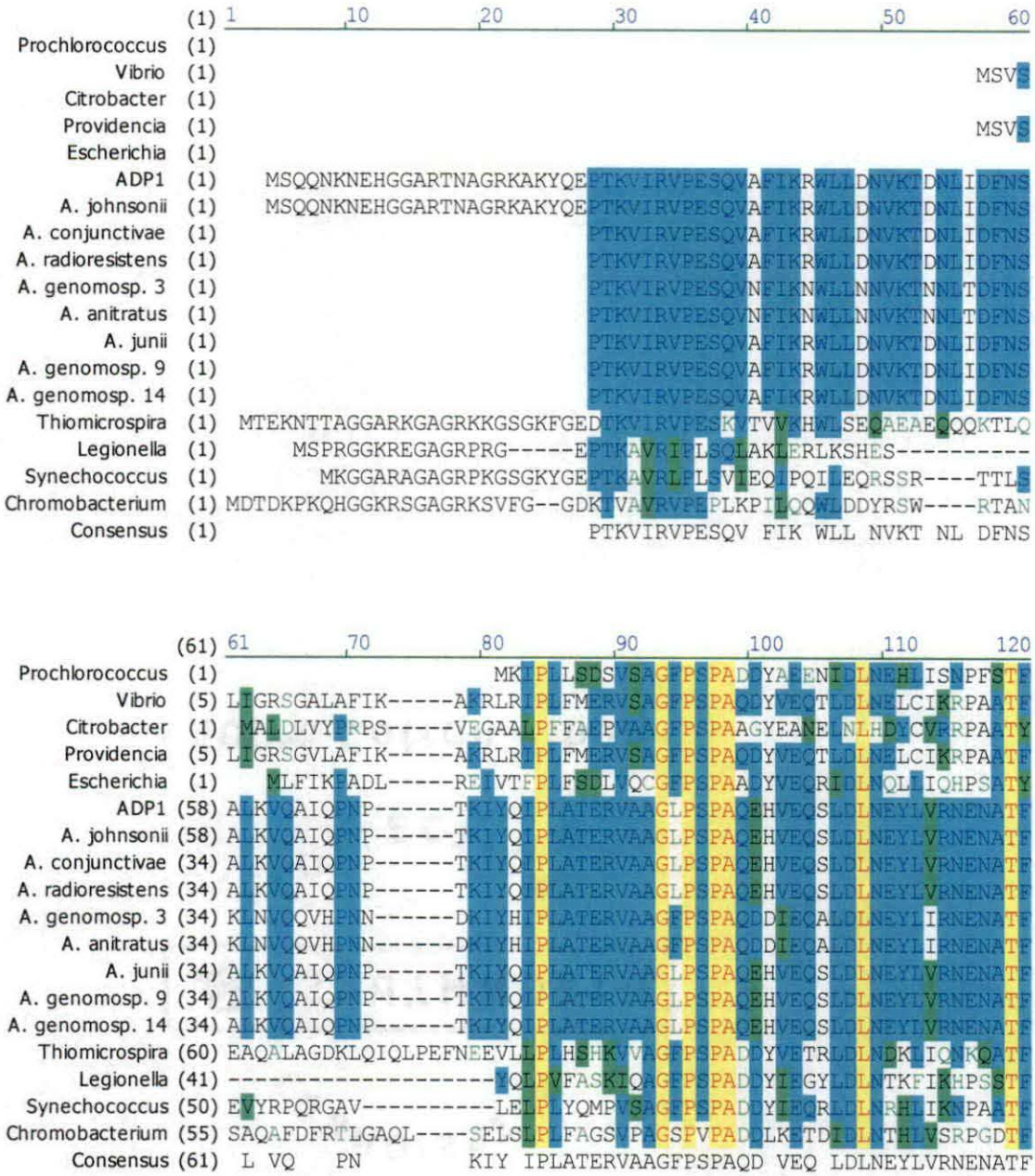


Figure 9. The multiple alignments of UmuD in different bacteria. This alignment includes the normal-sized UmuD homologs and the extra-long UmuD homologs in diverse bacteria.

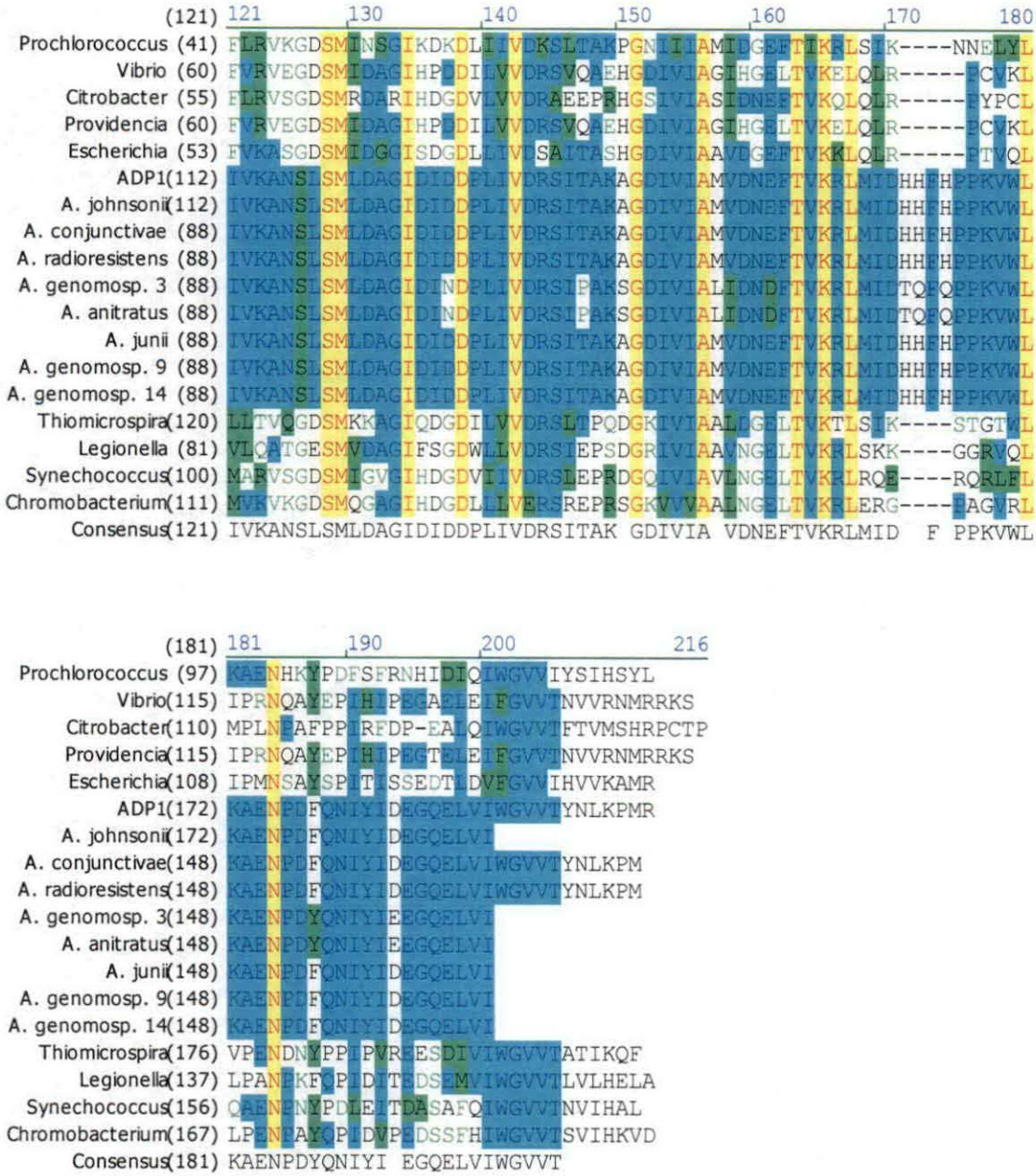


Figure 9. The multiple alignments of UmuD in different bacteria. (continued) This alignment includes the normal-sized UmuD homologs and the extra-long UmuD homologs in diverse bacteria.

The RecA-mediated self-cleavage site was changed to A-G in all aligned strains except for *E. coli*. However, the AG self-cleavage site is still functional, as it has been found in other active UmuD proteins in other model systems.

The catalytic residues, serine-60 and lysine-97, were all conserved in all bacteria. However, the leucine-101 and arginine-102 (LR motif), which are required for efficient self-cleavage of UmuD in *E. coli* (Sutton *et al.*, 2001) were only conserved in the normal-sized UmuD proteins, but not conserved in extra-long UmuD proteins, including ADP1 and examined *Acinetobacter* strains. In *Acinetobacter* species, this motif was changed to isoleucine-163 and aspartate-164 (Hare *et al.*, 2006). The additional five amino acids originally found uniquely in ADP1 were also found in the examined *Acinetobacter* strains, even though 2 strains possessed different residues.

Discussion

There are so many natural and artificial factors such as UV radiation, ionizing radiation, and chemical compounds that can cause various types of DNA damage, including base deletion/addition/substitution, DNA breakage, and deformations such as pyrimidine dimers. The outcome of these DNA damages may give rise to the malfunction of certain crucial proteins, and sometimes the DNA damage may interrupt the DNA replication in cells, causing cell death. Since DNA is crucial for life, DNA repair systems are required for all cells to recover from this damage. Error-free repair systems include both constitutive and inducible repair mechanisms, including proofreading repair, direct repair, excision repair, and can reverse or repair DNA damage to its original condition. On the other hand, cells may incur a cost of mutations when they apply the SOS response, an inducible error-prone system (Walker, 1996).

Even though the inducible error-prone repair system, SOS mutagenesis, can rescue the cells that are suffering a massive and lethal DNA damages, it is the last resort for cells to apply. Although this mechanism forces DNA replication, the lesion-bypass DNA polymerase often adds the wrong nucleotides opposed to a lesion sites such as abasic sites, UV-induced thymine dimers, and photoproducts, increasing the mutation frequency. Our understanding of the typical bacterial SOS mutagenesis system has been constructed based on experiments in *E. coli*, pioneered by Miroslav Radman in 1974 and developed by the Walker lab. Within this system, the

transcriptional and post-translational regulations are both necessary, including the inhibition and activation of SOS regulon genes and the modification of functional proteins.

Under normal conditions, the transcriptional repressor LexA protein binds tightly to the SOS box in the promoters of SOS genes. When DNA damage occurs, a SOS signal such as single-stranded DNA interacts with RecA, forming an active form RecA coprotease that inactivates LexA and thus activates SOS regulon expression. As a consequence, the many copies of SOS protein are encoded, including UmuD and UmuC. LexA and RecA themselves are SOS proteins, but the effectors in SOS mutagenesis are UmuD and UmuC, which are the precursors of DNA polymerase V ($UmuD'_2C$ complex).

The formation of functional DNA polymerase V is through a RecA-mediated self-cleavage, which can transform an inactive UmuD homodimer to an active UmuD' homodimer. The intact UmuD protein has two potential functions. Before the post-translational modification of the UmuD homodimer occurs, the intact UmuD protein functions as a cell cycle checkpoint inhibitor, providing cells more time to repair accurately (Opperman *et al.*, 1999). The other function of the intact UmuD protein is to shut off the SOS response by trapping UmuD' in the form of UmuD/UmuD' heterodimer (Battista *et al.*, 1990). The RecA-mediated self-cleavage site is found in the cysteine-24 and glycine-25 at amino-terminal tail of UmuD. In contrast to DNA polymerase III, DNA polymerase V is a low-replication-fidelity DNA polymerase because it does not have proofreading function. It explains why when translesion

DNA synthesis occurs, the mutation frequency increases (Banerjee *et al.*, 1990). Since the SOS mutagenesis is the last resort for cells, the intracellular amount of UmuD and UmuD' are under tight control by the ATP- dependent serine proteases, Lon and ClpXP. In the amino-terminus of UmuD proteins, there are several recognition sites for these proteases to recognize and then degrade UmuD and UmuD'.

This "typical" SOS mutagenesis model is also present in many other distant related bacteria, including the plant pathogen *Xanthomonas* and gram-positive bacteria *Mycobacterium and Bacillus*. Although they all have LexA and SOS boxes and require RecA, some variations are also found, such as the LexA gene numbers, SOS box nucleotide sequences and the actual activated SOS genes.

However, in some bacteria, the typical SOS mutagenesis system does not seem to be completely conserved including *Acinetobacter baylyi* strain ADP1 (Hare *et al.*, 2006; Rauch *et al.*, 1996). In contrast to *E. coli*, no LexA homolog and SOS box are found in ADP1 genome (Hare *et al.*, 2006). Moreover, the induction of *recA* is independent of the RecA protein (Rauch *et al.*, 1996). The *umuD* gene itself is longer in ADP1 than in *E. coli*, with an additional encoded 58 amino acids at its amino-terminus. In contrast to extra-long *umuD*, the *umuC* is mutated: there is no intact *umuC* found in ADP1.

In this project, eleven *Acinetobacter* strains were examined by the Touchdown PCR amplification to see whether or not the unusual *umuDC* features found in ADP1 are also present in other *Acinetobacter* species (see Table 4 for the overall summary).

Table 4. Overall summary for the analyses in the examined *Acinetobacter* strains.

Strain	Analyzed Seq. (bp)	Compared to ADP1 (nt)	Compared to ADP1 (a.a.)	Features
Acinetobacter junii	504	Identities = 504/504 (100%) Gaps = 0/504 (0%)	Identities = 167/167 (100%)	Extra-long N-terminus of UmuD Contain AG cleavage-site Contain Ser60/Lys97
Acinetobacter genomospecies 14	504	Identities = 504/504 (100%) Gaps = 0/504 (0%)	Identities = 167/167 (100%)	Extra-long N-terminus of UmuD Contain AG cleavage-site Contain Ser60/Lys97
Acinetobacter genomospecies 9	504	Identities = 504/504 (100%) Gaps = 0/504 (0%)	Identities = 167/167 (100%)	Extra-long N-terminus of UmuD Contain AG cleavage-site Contain Ser60/Lys97
Acinetobacter anitratus	504	Identities = 381/504 (76%) Gaps = 0/504 (0%)	Identities = 137/167 (82%) Similarities = 151/167 (90%) Gaps = 0/167 (0%)	Extra-long N-terminus of UmuD Contain AG cleavage-site Contain Ser60/Lys97
Acinetobacter genomospecies 3	504	Identities = 381/504 (76%) Gaps = 0/504 (0%)	Identities = 137/167 (82%) Similarities = 151/167 (90%) Gaps = 0/167 (0%)	Extra-long N-terminus of UmuD Contain AG cleavage-site Contain Ser60/Lys97
Acinetobacter radioresistens	651	Identities = 651/651 (100%) Gaps = 0/651 (0%)	UmuD homolog: Identities = 178/178 (100%) --- UmuC homolog: Identities = 35/35 (100%)	Extra-long N-terminus of UmuD Contain AG cleavage-site Contain Ser60/Lys97 Partial *umuC* region
Acinetobacter conjunctivae TU14	650	Identities = 649/651 (99%) Gaps = 1/651 (0%)	UmuD homolog: Identities = 178/178 (100%) --- UmuC homolog: Identities = 25/35 (71%)	Extra-long N-terminus of UmuD Contain AG cleavage-site Contain Ser60/Lys97 Partial *umuC* region
Acinetobacter johnsonii	770	Identities = 768/770 (99%) Gaps = 0/770 (0%)	UmuD homolog: Identities = 195/196 (99%) Similarities = 196/196 (100%)	Extra-long N-terminus of UmuD Contain AG cleavage-site Contain Ser60/Lys97 Contain *ddrR* homolog Contain *ddrR-umuD* promoter region: contain potential -35 and -10 promoters, and ribosome binding sites

These eleven strains are distributed throughout to the phylogenetic tree as determined by their 16S rRNA homology (Vaneechoutte *et al.*, 2006).

umuD homologs were present in eight out of eleven test strains. They all have a highly conserved (six have 100% identity; two have 82% identity) extra-long UmuD, implying that other strains in the genus may also possess the same unusual feature, whether at 100% or slightly less high level of conservation. Recently, however, *A. baumanii* strain ATCC 17978 has been observed to have a shorter UmuD (164 amino acids) which is more similar to the *E. coli* UmuD (Smith *et al.*, 2007).

The protein motif at the site leucine-101 and arginine-102 which is required for efficient self-cleavage of UmuD in *E. coli* (Sutton *et al.*, 2001) is changed to isoleucine-163 and aspartate-164 in ADP1 (Hare *et al.*, 2006). The changed protein, ID motif, was also present in these eight test strains. One interesting thing is that the LR motif is highly conserved in normal sized UmuD proteins, while it is changed in all of the extra-long UmuD proteins (Figure 9). This might indicate functional differences in requirements for cleavage of these UmuD homologs.

Moreover, the presence of additional five amino acids adjacent to aspartate-164 also present in these eight strains, even though the amino acid residues in *A. anitratus* and *A.* genomospecies 3 were different from those in other strains (Figure 9). However, these additional five amino acids may not to affect the function of UmuD because these residues do not interrupt the possible functional protein motifs and fall in a solvent-exposed loop according to the crystal structure of the *E. coli* UmuD' homodimer (Peat *et al.*, 1996; Ferentz *et al.*, 1997).

In addition to the *umuD* homologs, "partial" *umuC* homologs were also present in two (*A. radioresistens* and *A. conjunctivae* TU14) out of eleven test strains. The partial *umuC* homolog in *A. radioresistens* was 100% identical to that region in ADP1; however, a gap was found roughly two-thirds of the way through the coding region of *umuC* in *A. conjunctivae* TU14. As a result, a nonsense mutation was formed and the encoded amino acid leucine was replaced by a stop codon. Even though the stop codon was formed, it possibly will not cause additional problems to *A. conjunctivae* TU14 because even in ADP1, only thirty-nine amino acids are encoded, and these thirty-nine amino acids are not likely to form a functional protein.

One question arose when considering the unusual *umuC* in ADP1: is the *umuC* homolog also mutated in two fragments in these two strains? So far, we don't know because the primers used were only able to amplify the 5' region of the small *umuC* fragment. No PCR products were yielded with many sets of primers which were used trying to amplify the downstream regions of *umuDC* operon (Tables 2 & 3). In addition, no PCR products were yielded by an inverse PCR amplification protocol, either (data not shown).

The entire *ddrR-umuD* region was conserved in *A. johnsonii*, even though several nucleotides changed in the promoter region (Figure 7). These changes are not likely to affect the potential promoters because they are not located in the predicted promoter regions, -10 and -35. This result also implied that *ddrR* may also be conserved in this genus.

In *E. coli*, the intracellular levels of UmuD and UmuD' are regulated by Lon and ClpXP, respectively. The recognition sites for Lon and ClpXP are located in amino-terminus of UmuD. The alignments showed that both Lon recognition sites were highly conserved, especially the most important residue, proline, is completely conserved in all examined bacteria (Figure 9). However, the ClpXP recognition-site was not conserved in any aligned bacteria but *E. coli*. This result arose a question: is the ClpXP recognition site unique in *E. coli*, as are SOS boxes in various species? If UmuD in *Acinetobacter* is shown to self-cleave upon DNA damage then site-directed mutagenesis experiments in either UmuD or the Lon homolog may be performed to see if these protein motifs are functional.

The analyses from the multiple alignments of UmuD homologs (Figure 9) and the conservation of protein motifs suggest that the *Acinetobacter* UmuD may able to self-cleave. The catalytic residues, Ser-60 and Lys-97, are conserved in both *E. coli* and *Acinetobacter* species without any gaps, and the RecA*-mediated self-cleavage site, Ala-24/Gly-25, is also conserved in *Acinetobacter* species. One interesting thing found was that these functional catalytic residues, Ser-60 and Lys-97, are conserved in all bacteria regardless of whether they possess a normal-sized UmuD protein or an extra-long UmuD protein (Figure 9). Moreover, these bacteria like all others studied to date besides *E. coli*, possess a Ala/Gly self-cleavage site rather than a Cys-Gly site.

Clearly, Western blot experiments need to be conducted on DNA damage induced and uninduced ADP1 populations to determine whether UmuD can self-cleave, and if the cleavage is a part of the *Acinetobacter* DNA damage response.

From the BLAST of the ADP1 genome with the *E. coli* LexA, the best match was to UmuD. However, there is again, only ~37% identity in their carboxyl-terminal amino acids. No similarity exists at the amino-terminus of these proteins. The role of the extra-long amino-terminus of UmuD, if any, in the damage response of this genus remains to be determined.

Literature Cited

Ames, BN. 1979. Identifying environmental chemicals causing mutations and cancer. *Science*, **204**: 587-593.

Banerjee SK, BC Roshan, WL Christopher, and JE LeClerc. 1988. Frequency and spectrum of mutations produced by a single *cis*-syn thymine-thymine cyclobutane dimer in a single-stranded vector. *Proc. Natl. Acad. Sci.*, **85**: 8141-8145.

Banerjee SK, A Borden, RB Christensen, JE LeClerc, and CW Lawrence. 1990. SOS-dependent replication past a single *trans*-syn T-T cyclobutane dimer gives a different mutation spectrum and increased error rate compared with replication past this lesion in uninduced cells. *J. Bacteriol.*, **172**: 2105-2112.

Barbe V, D Vallenet, N Fonknechten, A Kreimeyer, S Oztas, L Labarre, S Cruveiller, C Robert, S Duprat, P Wincker, LN Ornston, J Weissenbach, P Marliere, GN Cohen, and C Medigue. 2004. Unique features revealed by the genome sequence of *Acinetobacter* sp. ADP1, a versatile and naturally transformation competent bacterium. *Nucleic Acids Res.* **32**: 5766-5779.

Battista JR, T Ohta, T Nohmi, W Sun, and GC Walker. 1990. Dominant negative *umuD* mutations decreasing RecA-mediated cleavage suggest roles for intact UmuD in modulation of SOS mutagenesis. *Proc. Natl. Acad. Sci.* **87**: 7190-7194.

Becherel OJ and RP Fuchs. 1999. SOS mutagenesis results from up-regulation of translesion synthesis. *J Mol Biol*, **294**: 299-306.

Berenstein D. 1987. UV-inducible DNA repair in *Acinetobacter calcoaceticus*. *Mutat. Res.* **183**: 219-224.

Brent R and M Ptashne. 1981. Mechanism of action of the *lexA* gene product. *Proc. Natl. Acad. Sci.*, **78**: 4204-4208.

Brooks PC, F Movahedzadeh, and EO Davis. 2001. Identification of some DNA damage-inducible genes of *Mycobacterium tuberculosis*: Apparent lack of correlation with LexA binding. *J. Bacteriol.*, **183**: 4459-4467.

Burckhardt SE, R Woodgate, RH Scheuermann, and H Echols. 1988. UmuD mutagenesis protein of *Escherichia coli*: Overproduction, purification, and cleavage by RecA. *Proc. Natl. Acad. Sci.*, **85**: 1811-1815.

Campoy S, G Mazon, AR Fernández de Henestrosa, M Llagostera, PB Monteiro, and J Barbé. 2002. A new regulatory DNA motif of the gamma subclass *Proteobacteria*: identification of the LexA protein binding site of the plant pathogen *Xylella fastidiosa*. *Microbiology* **148**: 3583-3597.

Carr EL, P Kämpfer, BKC Patel, V Gürtler and RJ Seviour. 2003. Seven novel species of *Acinetobacter* isolated from activated sludge. *Int J Syst Evol Microbiol.* **53**: 953-963.

Cheo DL, KW Bayles, and RE Yasbin. 1991. Cloning and characterization of DNA damage-inducible promoter regions from *Bacillus subtilis*. *J. Bacteriol.*, **173**: 1696-1703.

Chien M, Morozova I, Shi S, Sheng H, Chen J, Gomez SM, Asamani G, Hill K, Nuara J, Feder M, Rineer J, Greenberg JJ, Steshenko V, Park SH, Zhao B, Teplitskaya E, Edwards JR, Pampou S, Georghiou A, Chou IC, Iannuccilli W, Ulz ME, Kim DH, Geringer-Sameth A, Goldsberry C, Morozov P, Fischer SG, Segal G, Qu X, Rzhetsky A, Zhang P, Cayanis E, De Jong PJ, Ju J, Kalachikov S, Shuman HA, and Russo JJ. 2004. The genomic sequence of the accidental pathogen *Legionella pneumophila*. *Science*, **305**: 1966-1968.

Cox MM. 1991. The RecA protein as a recombinational repair system. *Mol. Microbiol.*, **5**: 1295-1299.

Cox EC. 1997. *mutS*, Proofreading and Cancer. *Genetics*, **146**: 443-446.

Craig NL and JW Roberts. 1981. Function of nucleoside triphosphate and polynucleotide in *Escherichia coli* recA protein-directed cleavage of phage lambda repressor. *J. Biol. Chem.*, **256**: 8039-8044.

Davis EO, EM Dullaghan, and L Rand. 2002. Definition of the mycobacterial SOS box and use to identify LexA-regulated genes in *Mycobacterium tuberculosis*. *J. Bacteriol.*, **184**: 3287-3295.

Don RH, PT Cox, BJ Wainwright, K Baker and JS Mattick. 1991. "Touchdown" PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Res.* **19**: 4008.

Feng, G, HC Tsui, and ME Winkler. 1996. Depletion of the cellular amounts of the MutS and MutH methyl-directed mismatch repair proteins in stationary-phase *Escherichia coli* K-12 cells. *J. Bacteriol.*, **178**: 2388-2396.

Ferentz AE, T Opperman, GC Walker, and G Wagner. 1997. Dimerization of the UmuD' protein in solution and its implications for regulation of SOS mutagenesis. *Nat. Struct. Biol.* **4**: 979-983.

Fernandez De Henestrosa AR, T Ogi, S Aoyagi, D Chafin, JJ Hayes, H Ohmori, and R Woodgate. 2000. Identification of additional genes belonging to the LexA regulon in *Escherichia coli: Mol Microbiol*, **35**: 1560-1568.

Fijalkowska IJ, RL Dunn, and RM Schaaper. 1997. Genetic requirements and mutational specificity of the *Escherichia coli* SOS mutator activity. *J. Bacteriol.*, **179**: 7435-7445.

Frank EG, DG Ennis, M Gonzalez, AS Levine, and R Woodgate. 1996. Regulation of SOS mutagenesis by proteolysis. *Proc. Natl. Acad. Sci.*, **93**: 10291-10296.

Friedberg EC, GC Walker, and W Siede. 1995. *DNA Repair and Mutagenesis.* ASM Press, Washington, D.C.

Fujiwara Y, C Masutani, T Mizukoshi, J Kondo, F Hanaoka, and S Iwai. 1999. Characterization of DNA recognition by the human UV-damaged DNA-binding protein. *J. Biol. Chem.*, **274**: 20027-20033.

Gonzalez M, EG Frank, AS Levine, and R Woodgate. 1998. Lon-mediated proteolysis of the *Escherichia coli* UmuD mutagenesis protein: in vitro degradation and identification of residues required for proteolysis. *Genes & Dev.* **12**: 3889-3999.

Gonzalez M, F Rasulova, MR Maurizi, and R Woodgate. 2000. Subunit-specific degradation of the UmuD/D' heterodimer by the ClpXP protease: the role of *trans* recognition in UmuD' stability. *EMBO J.* **19**: 5251-5258.

Goodman MF and R Woodgate. 2000. The biochemical basis and *in vivo* regulation of SOS-induced mutagenesis promoted by *Escherichia coli* DNA polymerase V ($UmuD'_2C$). *Cold Spring Harb Symp Quant Biol.* **65**: 31-40.

Hare JM, SN Perkins, and LA Gregg-Jolly. 2006. A constitutively expressed, truncated *umuDC* operon regulates the *recA*-dependent DNA damage induction of a gene in *Acinetobacter baylyi* strain ADP1. *Appl. Envir. Microbiol.*, **72**: 4036-4043.

Hartl DL, and EW Jones. 1999. *Essential Genetics,* 2nd edition. Jones and Bartlett Publishers, Sudbury, MA.

Higashitani N, A Higashitani, A Roth, and K Horiuchi. 1992. SOS induction in *Escherichia coli* by infection with mutant filamentous phage that are defective in initiation of complementary-strand DNA synthesis. *J. Bacteriol.*, **174:** 1612-1618.

Hübscher U, G Maga, and S Spadari. 2002. Eukaryotic DNA polymerases, *Annual Review of Biochemistry*, **71**: 133-163.

Jannasch HW and MJ Mottl. 1985. Geomicrobiology of deep-sea hydrothermal vents. *Science,* **229**: 717-725.

Jara M, C Núñuz, S Campoy, AR Fernández de Henestrosa, DR Lovley, and J Barbé. 2003. *Geobacter sulfurreducens* has two autoregulated *lexA* genes whose products do not bind the *recA* promoter: differing responses of *lexA* and *recA* to DNA damage. *J. Bacteriol.* **185**: 2493-2502.

Juni E. 1978. Genetics and physiology of *Acinetobacter*. *Annual Rev. Microbiol.* **32**: 349-371.

Khil PP and RD Camerini-Otero. 2002. Over 1000 genes are involved in the DNA damage response of *Escherichia coli. Mol. Microbiol.* **44**: 89-105.

Kitagawa Y, E Akaboshi, H Shinagawa, T Horii, H Ogawa, and T Kato. 1985. Structural analysis of the *umu* operon required for inducible mutagenesis in *Escherichia coli. Proc. Natl. Acad. Sci.* **82**: 4336-4340.

Lee MH, T Ohta, and GC Walker. 1994. A monocysteine approach for probing the structure and interactions of the UmuD protein. *J. Bacteriol.*, **176**: 4825-4837.

Little JW, SH Edmiston, LZ Pacelli, and DW Mount. 1980. Cleavage of the *Escherichia coli* LexA protein by the RecA protease. *Proc. Natl. Acad. Sci.* **77**: 3225-3229.

Little JW, DW Mount, and CR Yanisch-Perron. 1981. Purified LexA protein is a repressor of the *recA* and *lexA* genes. *Proc. Natl. Acad. Sci.*, **78**: 4199-4203.

Little JW and DW Mount. 1982. The SOS regulatory system of *Escherichia coli. Cell* **29**: 11-22.

Little JW. 1984. Autodigestion of LexA and phage λ repressors. *Proc. Natl. Acad. Sci.*, **81:** 1375-1379.

Lodish H, A Berk, SL Zipursky, P Matsudaira, D Baltimore, and J Darnell. 2001. *Molecular Cell Biology*, 4th edition. W. H. Freedman and Company, New York, NY.

Maki H. 2002. Origins of spontaneous mutations: specificity and directionality of base-substitution, frameshift, and sequence-substitution mutageneses. *Annual Review of Genetics,* **36**: 279-303.

Maor-Shoshani A, NB Reuven, G Tomer, and Z Livneh. 2000. Highly mutagenic replication by DNA polymerase V (UmuC) provides a mechanistic basis for SOS untargeted mutagenesis. *Proc. Natl. Acad. Sci.*, **97**: 565-570.

Masutani C, R Kusumoto, A Yamada, N Dohmae, M Yokoi, M Yuasa, M Araki, S Iwai, K Takio, and F Hanaoka. 1999. The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase eta. *Nature* **399**: 700-704.

Masutani C, M Araki, A Yamada, R Kusumoto, T Nogimori, T Maekawa, S Iwai, and F Hanaoka. 1999. Xeroderma pigmentosum variant (XP-V) correcting

protein from HeLa cells has a thymine dimer bypass DNA polymerase activity. *EMBO J,* **18**: 3491-3501.

McDonald JP, EG Frank, AS Levine, and R Woodgate. 1998. Intermolecular cleavage by UmuD-like mutagenesis proteins. *Proc. Natl. Acad. Sci.* **95**:1478-1483.

McNally KP, NE Freitag, and GC Walker. 1990. LexA-independent expression of a mutant *mucAB* operon. *J. Bacteriol.*, **172**: 6223-6231.

Metzgar D, and JM Bacher. 2004. *Acinetobacter* sp. ADP1: an ideal model organism for genetic analysis and genome engineering. *Nucleic Acids Res.* **32**: 5780-5790.

Mount DW, KB Low, and S Edmiston. 1972. Dominant mutations (*lex*) in *Escherichia coli* K12 which affect radiation sensitivity and frequency of ultraviolet light-induced mutations. *J. Bacteriol.* **112**: 886-893.

Movahedzadeh F, MJ Colston, and EO Davis. 1997. Characterization of *Mycobacterium tuberculosis* LexA: recognition of a Cheo (Bacillus-type SOS) box. *Microbiology*, **143**: 929-936.

Nemec A, L Dijkshoorn, and P Ješek. 2000. Recognition of two novel phenons of the genus *Acinetobacter* among glucose non-acidifying isolates from human specimens. *J. Clin. Microbiol.* **38**: 3937-3941.

Nemec A, De Baere T, I Tjernberg, M Vaneechoutte, TJK van der Reijden, and L Dijkshoorn. 2001. *Acinetobacter ursingii* sp. nov. and *Acinetobacter schindleri* sp. nov., isolated from human clinical specimens. *Int J Syst Evol Microbiol,* **51**: 1891-1899.

Nemec A, L Dijkshoorn, I Cleenwerck, T De Baere, D Janssens, TJK van der Reijden, P Jezek, and M Vaneechoutte. 2003. *Acinetobacter parvus* sp. nov., a small-colony-forming species isolated from human clinical specimens. *Int J Syst Evol Microbiol*, **53**: 1563-1567.

Nohmi T, JR Battista, LA Dodson, and GC Walker. 1988. RecA-mediated cleavage activates UmuD for mutagenesis: mechanistic relationship between

transcriptional derepression and posttranslational activation. *Proc. Natl. Acad. Sci.* **85**: 1816-1820.

Ohta T, MD Sutton, A Guzzo, S Cole, AE Ferentz, and GC Walker. 1999. Mutations affecting the ability of the *Escherichia coli* UmuD' protein to participate in SOS mutagenesis. *J. Bacteriol.*, **181**: 177-185.

Opperman T, S Murli, BT Smith, and GC Walker. 1999. A model for a *umuDC*-dependent prokaryotic DNA damage checkpoint. *Proc. Natl. Acad. Sci.*, **96**: 9218-9223.

Peat TS, EG Frank, JP McDonald, AS Levine, R Woodgate, and WA Hendrickson. 1996. Structure of the UmuD' protein and its regulation in response to DNA damage. *Nature*, **380**: 727-730.

Peat, TS, EG Frank, JP McDonald, AS Levine, R Woodgate, and WA Hendrickson. 1996. The UmuD' protein filament and its potential role in damage induced mutagenesis. *Structure*, **4**: 1401-1412.

Perry KL, SJ Elledge, BB Mitchell, L Marsh, and GC Walker. 1985. *umuDC* and *mucAB* operons whose products are required for UV light- and chemical-induced mutagenesis: UmuD, MucA, and LexA proteins share homology. *Proc. Natl. Acad. Sci.*, **82:** 4331-4335.

Radman, M. 1974. Phenomenology of an inducible mutagenic DNA repair pathway in *Escherichia coli*: SOS repair hypothesis, p. 128-142. *In* L. Prokash, F. Sherman, M. Miller, C. Lawrence, and H. W. Tabor (ed.), *Molecular and environmental aspects of mutagenesis*. Charles C Thomas Publisher, Springfield, Ill.

Rajagopalan M, C Lu, R Woodgate, M O'Donnell, MF Goodman, and H Echols. 1992. Activity of the purified mutagenesis proteins UmuC, UmuD, and RecA in replicative bypass of an abasic DNA lesion by DNA polymerase III. *Proc. Natl. Acad. Sci.*, **89**: 10777-10781.

Rauch PJ, R Palmen, AA Burds, LA Gregg-Jolly, JR van der Zee, and KJ Hellingwerf. 1996. The expression of the *Acinetobacter calcoaceticus recA* gene

increases in response to DNA damage independently of RecA and of development of competence for natural transformation. *Microbiology* **142**: 1025-1032.

Reuven NB, G Tomer, and Z Livneh. 1998. The mutagenesis proteins UmuD' and UmuC prevent lethal frameshifts while increasing base substitution mutations. *Mol. Cell,* **2:** 191-199.

Reuven NB, G Arad, A Maor-Shoshani, and Z Livneh. 1999. The mutagenesis protein UmuC is a DNA polymerase activated by UmuD', RecA, and SSB and is specialized for translesion replication. *J. Biol. Chem.*, **274**: 31763-31766.

Roux, KH. 1994. Using mismatched primer-template pairs in Touchdown PCR. *BioTechniques*, **16**: 812-814.

Sassanfar M, and JW Roberts. 1990. Nature of the SOS-inducing signal in E*scherichia coli*. The involvement of DNA replication. *J Mol Biol*, **212**: 79-96.

Schnarr M, J Pouyet, M Granger-Schnarr, and M Daune. 1985. Large-scale purification, oligomerization equilibria, and specific interaction of the LexA repressor of *Escherichia coli. Biochemistry*, **24**: 2812-2818.

Scott KM, Sievert SM, Abril FN, Ball LA, Barrett CJ, Blake RA, Boller AJ, Chain PS, Clark JA, Davis CR, Detter C, Do KF, Dobrinski KP, Faza BI, Fitzpatrick KA, Freyermuth SK, Harmer TL, Hauser LJ, Hugler M, Kerfeld CA, Klotz MG, Kong WW, Land M, Lapidus A, Larimer FW, Longo DL, Lucas S, Malfatti SA, Massey SE, Martin DD, McCuddin Z, Meyer F, Moore JL, Ocampo LH Jr, Paul JH, Paulsen IT, Reep DK, Ren Q, Ross RL, Sato PY, Thomas P, Tinkham LE, Zeruth GT. 2006. The genome of deep-sea vent chemolithoautotroph *Thiomicrospira crunogena* XCL-2. *PLoS Biol.* **4**: e383.

Singer B, and JT Kusmierek. 1982. Chemical mutagenesis. *Annu Rev Biochem*, **51**: 655-693.

Shinagawa H, H Iwasaki, T Kato, and A Nakata. 1988. RecA protein-dependent cleavage of UmuD protein and SOS mutagenesis. *Proc. Natl. Acad. Sci.*, **85**: 1806-1810.

Smith CM, WH Koch, SB Franklin, PL Foster, TA Cebula, and E Eisenstadt. 1990. Sequence analysis and mapping of the *Salmonella typhimurium* LT2 *umuDC* operon. *Journal of Bacteriology,* **172**: 4964-4978.

Smith BT, and GC Walker. 1998. Mutagenesis and more: *umuDC* and the *Escherichia coli* SOS response. *Genetics,* **148**: 1599-1610.

Smith CK, TA Baker, and RT Sauer. 1999. Lon and Clp family proteases and chaperones share homologous substrate-recognition domains. *Proc. Natl. Acad. Sci.*, **96**: 6678-6682.

Sommer S, J Knezevic, A Bailone, and R Devoret. 1993. Induction of only one SOS operon, *umuDC*, is required for SOS mutagenesis in *Escherichia coli. Mol Gen Genet*, **239**: 137-144.

Story RM, DK Bishop, N Kleckner, and TA Steitz. 1993. Structural relationship of bacterial RecA proteins to recombination proteins from bacteriophage T4 and yeast. *Science*, **259**: 1892-1896.

Sugita C, Ogata K, Shikata M, Jikuya H, Takano J, Furumichi M, Kanehisa M, Omata T, Sugiura M, Sugita M. 2007. Complete nucleotide sequence of the freshwater unicellular cyanobacterium *Synechococcus elongatus* PCC 6301 chromosome: gene content and organization. *Photosynth Res.* [Epub ahead of print]

Sutton MD, T Opperman, and GC Walker. 1999. The *Escherichia coli* SOS mutagenesis proteins UmuD and UmuD' interact physically with the replicative DNA polymerase. *Proc. Natl. Acad. Sci.*, **96**: 12373-12378.

Sutton, MD, BT Smith, VG Godoy, and GC Walker. 2000. The SOS response: recent insights into *umuDC*-dependent mutagenesis and DNA damage tolerance. *Annu Rev Genet*, **34**: 479-482.

Sutton MD, M Kim, and GC Walker. 2001. Genetic and biochemical characterization of a novel *umuD* mutation: Insights into a mechanism for UmuD self-cleavage. *Journal of Bacteriology*, **183**: 347-350.

Tang MG, I Bruck, R Eritja, J Turner, EG Frank, R Woodgate, M O'Donnell, and MF Goodman. 1998. Biochemical basis of SOS-induced mutagenesis in *Escherichia coli*: Reconstitution of *in vitro* lesion bypass dependent on the UmuD'$_2$C mutagenic complex and RecA protein. *Proc. Natl. Acad. Sci.*, **95**: 9755-9760.

Tang MG, X Shen, EG Frank, M O'Donnell, R Woodgate, and MF Goodman. 1999. UmuD'$_2$C is an error-prone DNA polymerase, *Escherichia coli* pol V. *Proc. Natl. Acad. Sci.*, **96**: 8919-8924.

Tang MG, P Pham, X Shen, JS Taylor, M O'Donnell, R Woodgate, and MF Goodman. 2000. Roles of *E. coli* DNA polymerases IV and V in lesion-targeted and untargeted SOS mutagenesis. *Nature*, **404**: 1014-1918.

Thliveris AT, JW Little, and DW Mount. 1991. Repression of the *E. coli recA* gene requires at least two LexA protein monomers. *Biochimie*, **73**: 449-456.

Thinh LV, and DJ Griffiths. 1973. Nucleic acid synthesis accompanying the recovery of cell division and chloroplast development in "giant" cells of the Emerson strain of *Chlorella. Plant Cell Physiol.*, **14**: 497-500.

Vasconcelos ATR, de Almeida,D.F., Almeida,F.C., de Almeida,L.G.P., de Almeida,R., Goncalves,J.A.A., Andrade,E.M.,Antonio,R.V., Araripe,J., de Araujo,M.F.F., Filho,S.A., Azevedo,V.,Batista,A.J., Bataus,L.A.M., Batista,J.S., Belo,A., vander Berg,C., Blamey,J., Bogo,M., Bonato,S., Bordignon,J., Brito,C.A., Brocchi,M., Burity,H.A., Camargo,A.A., Cardoso,D.D.P., Carneiro,N.P., Carraro,D.M., Carvalho,C.M.B., Cascardo,J.C.M., Cavada,B.S., Chueire,L.M.O., Pasa,T.B.C., Duran,N., Fagundes,N., Falcao,C.L., Fantinatti,F., Farias,I.P., Felipe,M.S.S., Ferrari,L.P., Ferro,J.A., Ferro,M.I.T., Franco,G.R., Freitas,N.S.A., Furlan,L.R., Gazzinelli,R.T., Gomes,E.A., Goncalves,P.R., Grangeiro,T.B., Grattapaglia,D., Grisard,E.C., Guimaraes,C.T., Hanna,E.S., Hungria,M., Jardim,S.N., Laurino,J., Leoi,L.C.T., Fassarella,L., Lima,A., Loureiro,M.F., Lyra,M.C.P., Macedo,M., Madeira,H.M.F., Manfio,G.P., Maranhao,A.Q., Martins,W.S., di Mauro,S.M.Z., de Medeiros,S.R.B., Meissner,R.D.V., Menck,C.F.M., Moreira,M.A.M., Nascimento,F.F., Nicolas,M.F., Oliveira,J.G., Oliveira,S.C., Paixao,R.F.C., Parente,J.A.,

Pedrosa,F.O., Pena,S.J.D., Perreira,J.O., Perreira,M., Pinto,L.S.R.C., Pinto,L.S., Porto,J.I.R., Potrich,D.P., Neto,C.E.R., Reis,A.M.M., Rigo,L.U., Rondinelli,E., dos Santos,E.B.P., Santos,F.R., Schneider,M.P.C., Seuanez,H.N., Silva,A.M.R., da Silva,A.L.C., Silva,D.W., Silva,R., Simoes,I.C., Simon,D., Soares,C.M.A., Soares,R.B.A., Souza,E.M., Souza,K.R.L., Souza,R.C., Steffens,M.B.R., Steindel,M., Teixeira,S.R., Urmenyi,T., Vettore,A., Wassem,R., Zaha,A. and Simpson,A.J.G. 2003. The complete genome sequence of *Chromobacterium violaceum* reveals remarkable and exploitable bacterial adaptability. *Proc. Natl. Acad. Sci.* **100**: 11660-11665.

Walker GC. 1984. Mutagenesis and inducible responses to deoxyribonucleic acid damage in *Escherichia coli*. *Microbiol. Mol. Biol. Rev.* **48**: 60-93

Walker GC. 1985. Inducible DNA repair systems. *Annu. Rev. Biochem.* **54**: 425-457.

Walker GC. 1996. The SOS response of *Escherichia coli*, p. 1400-1416. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli and Salmonella: cellular and molecular biology.* ASM Press, Washington, D.C.

Watson J, and F Crick. 1953. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature,* **171**: 737-738.

Whitworth GB, and LA Gregg-Jolly. 2000. The regulation of a DNA damage-inducible locus in *Acinetobacter* sp. strain ADP1. *Abstracts of the Annual Meeting of the American Society for Microbiology*, H-60.

Winterling KW, D Chafin, JJ Hayes, J Sun, AS Levine, RE Yasbin, and R Woodgate. 1998. The *Bacillus subtilis* DinR binding site: redefinition of the consensus sequence. *J. Bacteriol.*, **180**: 2201-2211.

Woodgate R, M Rajagopalan, C Lu, and H Echols. 1989. UmuC mutagenesis protein of *Escherichia coli*: Purification and interaction with UmuD and UmuD'. *Proc. Natl. Acad. Sci.*, **86**: 7301-7305.

Yang YC, CH Hsu, CP Chou, and MK Yang. 2002. Genetic organization of the *lexA*, *recA* and *recX* genes in *Xanthomonas campestris*. *Microbiol Lett*, **209**: 149-154.

Yang MK, SR Su, and VL Sung. 2005. Identification and characterization of a second *lexA* gene of *Xanthomonas axonopodis* pathovar citri. *Appl. Envir. Microbiol.*, **71**: 3589-3598.

Yasuda T, K Morimatsu, T Horii, T Nagata, and H Ohmori. 1998. Inhibition of *Escherichia coli* RecA coprotease activities by DinI. *EMBO J*, **17**: 3207-3213.